BULLETIN

OF THE

AMERICAN MATHEMATICAL SOCIETY

EDITED BY

EARLE R. HEDRICK

DAVID R. CURTISS WILLIAM R. LONGLEY

WITH THE ASSISTANCE OF

HAROLD T. DAVIS HARRY W. KUHN

THEOPHIL H. HILDEBRANDT DERRICK N. LEHMER

JOHN R. KLINE HENRY L. RIETZ

CAROLINE E. SEELY

VOLUME XXXIX, NUMBER 10

OCTOBER, 1933

PUBLISHED BY THE SOCIETY

MENASHA, WIS., AND NEW YORK

1933

PUBLISHED MONTHLY

Whole No. 409 $9.00 a Year

Entered as second class matter April 24, 1926 at the post office at Menasha, Wis., under the act of August 24, 1912.

AMERICAN MATHEMATICAL SOCIETY

Mail Address: 501 West 116th Street, New York, N.Y., for the Secretary, Treasurer, Librarian, and Clerk.

Office: Room 413, Physics Building, Columbia University, New York City.

President: Professor A. B. Coble, University of Illinois, Urbana, Ill.

Secretary: Dean R. G. D. Richardson, Brown University, Providence, R.I.

Associate Secretaries: Professor M. H. Ingraham, University of Wisconsin, Madison, Wis.; Professor J. R. Kline, University of Pennsylvania, Philadelphia, Pa.; Dean T. M. Putnam, University of California, Berkeley, Calif.

Treasurer: Professor G. W. Mullins, Barnard College in Columbia University, New York City.

Librarian: Professor R. C. Archibald, Brown University, Providence, R.I.

Clerk: Dr. Caroline E. Seely, 413 Physics Building, Columbia University, New York City.

Editor-in-Chief of the Bulletin: Professor E. R. Hedrick, University of California at Los Angeles, Los Angeles, Calif.

Managing Editor of the Transactions: Professor R. D. Carmichael, University of Illinois, Urbana, Ill.

Chairman, Colloquium Editorial Committee: Professor R. L. Moore, University of Texas, Austin, Tex.

Managing Editor of the American Journal of Mathematics: Professor Abraham Cohen, Johns Hopkins University, Baltimore, Md.

Representatives of the Society in the Division of Physical Sciences of the National Research Council: Professors A. B. Coble, E. R. Hedrick, and Dean R. G. D. Richardson.

Representatives on the Council of the American Association for the Advancement of Science: Professors Louis Ingold and M. H. Ingraham.

Chairman, Committee on Arrangements for the Semi-Centennial: Professor T. S. Fiske, Columbia University, New York City.

QUANTUM MECHANICS AND ASYMPTOTIC SERIES†

BY G. D. BIRKHOFF

PART I.

1. *Introduction.* In its bold primary outline the program of quantum mechanics in the Schrödinger form runs as follows:

(A) Set up the Hamiltonian equations of the atomic system (nucleus+electrons) on a classical basis:

$$\frac{dx_i}{dt} = \frac{\partial H}{\partial y_i}; \quad \frac{dy_i}{dt} = -\frac{\partial H}{\partial x_i}, \qquad (i = 1, \cdots, n),$$

where x_i, y_i are ordinary rectangular coordinates and momenta, and $H(x_1, \cdots, x_n; y_1, \cdots, y_n)$ is the total energy. The associated Hamilton-Jacobi partial differential equation is then

$$\frac{\partial S}{\partial t} + H\left(x_1, \cdots, x_n; \frac{\partial S}{\partial x_1}, \cdots, \frac{\partial S}{\partial x_n}\right) = 0.$$

(B) Write down the corresponding homogeneous linear partial differential equation (the Schrödinger wave equation):

$$\frac{1}{\lambda}\frac{\partial \psi}{\partial t} + H\left(x_1, \cdots, x_n; \frac{1}{\lambda}\frac{\partial}{\partial x_1}, \cdots, \frac{1}{\lambda}\frac{\partial}{\partial x_n}\right)\psi = 0,$$

where the operational symbols in H appear on the right hand side of the individual terms of H, and where $\lambda = 2\pi i/h$, if h is Planck's constant.

(C) Write

$$\psi = e^{-\lambda E t}\psi^*(x_1, \cdots, x_n),$$

thus obtaining the linear differential equation (written in operational form)

$$(H - E)\psi^* = 0,$$

and determine the characteristic solutions, $\psi_1^*, \psi_2^*, \cdots$, for which the ψ_i^* vanish at infinity in such wise that

† An address delivered at Chicago, June 20, 1933, before the Society and Section A of the American Association for the Advancement of Science. In connection with the first part of this lecture, see two notes in the Proceedings of the National Academy of Sciences for March and April, 1933.

$$\int |\psi_i^*|^2 dx_1 \cdots dx_n$$

is finite. The corresponding E_1, E_2, $\cdots$ then prescribe the possible "energy levels," so that the possible "spectral frequencies," ν_{mn}, are those given by the formula

$$h\nu_{mn} = E_m - E_n, \qquad (E_m > E_n),$$

in accordance with the Planck-Einstein law.

Thus the program begins by relating the physical problem to a special linear boundary value problem of classical type. In its further development the aim is to obtain a complete account of atomic properties which is in accord with this starting point. While there has been extraordinary progress, there can be little doubt that complete success of the program is hardly to be hoped for.

My purpose today is to lay before you a tentative answer to two important mathematical questions raised by the primary program itself.

Firstly, what is the *mathematical* significance of the Schrödinger wave equation in its relation to the Hamiltonian equations? My answer will be given in terms of classic formal processes connected with asymptotic series. It is true that the theoretical physicist has obtained a kind of "deduction" of the wave equation on the basis of the analogy between the wave theory of light and the elementary optical theory (wave and particle theory). But I hope to bring out more clearly the true inwardness of the Schrödinger wave equation as a purely mathematical entity.†

Secondly, the form of the wave equation which arises in practice is usually reducible by means of separation of variables to an ordinary differential equation essentially of the following type:

$$\frac{d^2\psi}{dx^2} + \frac{8\pi^2 m}{h^2}(E - V(x))\psi = 0,$$

where the function $V(x)$ is defined in a certain interval in which $E - V(x)$ changes sign. For the determination of the character-

† For references to the important earlier work of de Broglie, L. Brillouin, Schrödinger, and Dirac, see my notes cited above.

istic numbers and functions, certain asymptotic series have been employed by Wentzel, Brillouin and Kramers.† It is, however, an open question as to whether or not their methods are justifiable. However, in an important recent paper bearing directly upon the points at issue, Langer‡ announces that he intends later to give a general discussion of this and similar questions. My intention here is to outline a simple justification of the Wentzel-Brillouin-Kramers method on the basis of a slight extension of some earlier results of my own.

Thus asymptotic series play a central role in what I have to say today. I would not be surprised if such series were found ultimately to be of importance in other aspects of quantum mechanics; for example, in connection with the proper formulation of Heisenberg's uncertainty principle.

2. *Linear Equations and Asymptotic Series.* Let

$$L(\psi, \lambda) = 0 \tag{1}$$

be any linear homogeneous differential equation in the dependent variable ψ and the independent variables $x_1, x_2, \cdots, x_n$, and involving λ, where λ is a large parameter. This equation will be ordinary or partial according as $n=1$ or $n>1$. Let us suppose that the coefficients of ψ and of its derivatives in $L(\psi, \lambda)$ are analytic in $x_1, \cdots, x_n$, and λ, and expansible in convergent power series in $1/\lambda$ for $|\lambda|>\Lambda$.

Now under these circumstances it has frequently been found that the differentiation of certain solutions ψ, as well as of their various derivatives with respect to x_i, is asymptotically equivalent to multiplication by $\lambda \partial S/\partial x_i$, where S is a suitable function of $x_1, \cdots, x_n$. For instance in the simple case $n=1$ of an equation

$$L(\psi, \lambda) \equiv \frac{1}{\lambda^2}\frac{d^2\psi}{dx^2} + \psi = 0 \tag{1a}$$

with solutions $e^{\pm \lambda i x}$, we have $d\psi/dx = \pm \lambda i \psi$, so that $S = \pm ix$ in

† For the principal references see a note by J. L. Dunham, *On Wentzel-Brillouin-Kramers' method of solving the wave equation*, Physical Review, vol. 41 (Sept. 15, 1932).

‡ R. E. Langer, *On the asymptotic solutions of differential equations, with an application to the Bessel functions of large complex order*, Transactions of this Society, vol. 34 (1932), pp. 447–480.

this case. Again, in the case of the Fourier's equation with independent variable taken as λx,

$$(1b) \qquad L(\psi, \lambda) \equiv \frac{1}{\lambda^2}\frac{d^2\psi}{dx^2} + \frac{1}{\lambda^2}\frac{1}{x}\frac{d\psi}{dx} + \psi = 0,$$

we have $S = \pm ix$ as before for suitable solutions.

In consequence of such an asymptotic relationship we are led to write

$$\psi \sim e^{\lambda S(x_1, \cdots, x_n)} v_0(x_1, \cdots, x_n)$$

as a first approximation to ψ, and thence successively to an asymptotic series for ψ:

$$\psi \sim e^{\lambda S}\left(v_0 + \frac{v_1}{\lambda} + \frac{v_2}{\lambda^2} + \cdots\right),$$

where $v_0, v_1, \cdots$ are functions of $x_1, \cdots, x_n$. The precise test for such an asymptotic series solution is of course that when it is substituted for ψ in $L(\psi, \lambda)$, with the indicated differentiations carried out and the coefficients of like powers of λ collected according to the usual formal rules, the expression $L(\psi, \lambda)$ reduces identically to 0. It is not to be expected in general that such a series converges and yields an actual solution, although this may occur in special cases. Obviously it may be assumed that v_0 does not vanish identically, since otherwise we could remove a factor $1/\lambda$ from the solution.

We shall term S a "phase function," and any corresponding v_0 an "amplitude function" of S. The reasons for this designation will appear subsequently.

We propose first to outline some fundamental facts concerning this classical formal process which have apparently escaped attention. For this purpose we find it convenient to introduce the modified differential operators

$$\frac{\partial^{[1]}F}{\partial x_i} = \frac{1}{\lambda}\frac{\partial F}{\partial x_i}, \quad \frac{\partial^{[2]}F}{\partial x_i \partial x_j} = \frac{1}{\lambda^2}\frac{\partial^2 F}{\partial x_i \partial x_j}, \text{ etc.}\dagger$$

In the two special cases noted above this notation allows us to write the equations (1a), (1b) as follows:

† See my paper, Transactions of this Society, vol. 9 (1908), pp. 219–231.

$$\frac{d^{[2]}\psi}{dx^2} + \psi = 0; \quad \frac{d^{[2]}\psi}{dx^2} + \frac{1}{\lambda}\frac{1}{x}\frac{d^{[1]}\psi}{dx} + \psi = 0,$$

while, more generally, $L(\psi, \lambda)$ may be written as a power series in $1/\lambda$,

$$L(\psi, \lambda) \equiv L_0(\psi) + \frac{1}{\lambda}L_1(\psi) + \cdots, \tag{2}$$

where $L_i(\psi)$, $(i = 0, 1, \cdots)$, are linear homogeneous expressions in ψ, $\partial^{[1]}\psi/\partial x_i$, $\partial^{[2]}\psi/\partial x_i \partial x_j$, etc., and where we may assume that $L_0(\psi) \not\equiv 0$, since we may always multiply through by a suitable power of λ. Evidently, then, we may write

$$L_i(\psi) = \xi^{(i)}\psi + \sum_j \xi_j^{(i)} \frac{\partial^{[1]}\psi}{\partial x_j} + \sum_{jk} \xi_{jk}^{(i)} \frac{\partial^{[2]}\psi}{\partial x_j \partial x_k} + \cdots. \tag{3}$$

Furthermore, on account of the interchangeability of the order of differentiation, we may assume

$$\xi_{jk}^{(i)} = \xi_{jk}^{(i)}; \quad \xi_{jkl}^{(i)} = \xi_{jlk}^{(i)} = \xi_{klj}^{(i)} = \cdots; \cdots. \tag{4}$$

The order of the expression $L(\psi, \lambda)$ is evidently that of the highest order of $L_0, L_1, \cdots$, say m. We shall assume that m is the actual order of $L_0(\psi)$, and shall term $L_0(\psi)$ the "principal part" of $L(\psi, \lambda)$; in dealing with L_0 we shall omit the superscripts in referring to $\xi^{(0)}, \xi_i^{(0)}, \cdots$.

Later on we shall have something to say concerning the existence of actual solutions of (1) corresponding to such asymptotic series solutions in the special case $m = 2$. For the present we shall make only the following heuristic remarks. (1) Solutions asymptotically represented by such series solutions will exist for suitably restricted ranges of the variables $x_1, \cdots, x_n$. (2) These actual solutions are not uniquely determined by the series solutions; for example, if ψ is so represented, so also will $(1+e^{-\lambda})\psi$ be if λ is real and positive. (3) The solution $c_1\psi_1 + c_2\psi_2$, where ψ_1 and ψ_2 are represented by two such series, is in general represented asymptotically by the dominant one of the series for $c_1\psi_1$ or $c_2\psi_2$.

If we substitute in the hypothetical series for ψ and equate the term independent of λ to zero, we obtain an equation

$$P\left(x_1, \cdots, x_n; \frac{\partial S}{\partial x_1}, \cdots, \frac{\partial S}{\partial x_n}\right) = 0 \tag{5}$$

on removal of a factor $e^{\lambda S}v_0$. Here the explicit expression for P is

$$P \equiv \xi + \sum_i \xi_i \frac{\partial S}{\partial x_i} + \sum_{ij} \xi_{ij} \frac{\partial S}{\partial x_i} \frac{\partial S}{\partial x_j} + \cdots . \tag{6}$$

Thus P is a polynomial of degree m in $\partial S/\partial x_i$, $(i = 1, \cdots, n)$, in which one or more terms of degree m are actually present.

We shall term (5) the "multiplier equation" for $L(\psi, \lambda) = 0$. This is an equation of the first order in S which does not contain S, and which is in general non-linear. In the two examples above the multiplier equation is

$$1 + \left(\frac{dS}{dx}\right)^2 = 0,$$

so that we find $S = \pm ix$, up to an additive constant.

It is also to be noted that for a given polynomial P there is one and only one corresponding principal part L_0, while L_1, L_2, $\cdots$ remain entirely arbitrary. For convenience we shall term the special case in which $L \equiv L_0$ the "principal equation" for the given multiplier equation.

Let us proceed to the determination of v_0, v_1, $\cdots$, which turn out to be respectively determined by the later equations for $k = 1, 2, \cdots$. We have then to substitute the expressions for ψ, $\partial^{[1]}\psi/\partial x_i$, $\cdots$ in the power series for $L(\psi, \lambda)$, and equate the coefficients of $1/\lambda$, $1/\lambda^2$, $\cdots$ to zero. In the case $k = 1$, we observe first that a single term only, $Qe^{\lambda S}v_0/\lambda$, is contributed by the terms after the first in the series for $L(\psi, \lambda)$; here

$$Q = \xi^{(1)} + \sum_i \xi_i^{(1)} \frac{\partial S}{\partial x_i} + \cdots .$$

Thus Q is related to L_1 just as P is to L_0. It remains then to determine the term in $1/\lambda$ which arises from L_0. To do so we note that, to terms of the first order in $1/\lambda$,

$$
\begin{aligned}
\frac{\partial^{[1]}\psi}{\partial x_i} &\sim e^{\lambda S}\left(\frac{\partial S}{\partial x_i}v_0 + \frac{1}{\lambda}\left(\frac{\partial v_0}{\partial x_i} + \frac{\partial S}{\partial x_i}v_1\right)\right) \\
&\sim e^{\lambda S}\left(\left[\frac{\partial S}{\partial x_i} + \frac{1}{\lambda}\frac{\partial}{\partial x_i}\right]v_0 + \frac{1}{\lambda}\frac{\partial S}{\partial x_i}v_1\right), \\
\frac{\partial^{[2]}\psi}{\partial x_i \partial x_j} &\sim e^{\lambda S}\left(\left[\frac{\partial S}{\partial x_i} + \frac{1}{\lambda}\frac{\partial}{\partial x_i}\right]\left[\frac{\partial S}{\partial x_j} + \frac{1}{\lambda}\frac{\partial}{\partial x_j}\right]v_0\right. \\
&\qquad \left. + \frac{1}{\lambda}\frac{\partial S}{\partial x_i}\frac{\partial S}{\partial x_j}v_1\right),
\end{aligned} \tag{7}
$$

and so on. Here we use an obvious operational notation.

Now we observe that on substitution in L_0, the terms in v_1 disappear identically because of the multiplier equation. Hence aside from the factor $e^{\lambda S}$, the coefficient of $1/\lambda$ in L_0 is

$$
\begin{aligned}
&\sum_i \xi_i \frac{\partial v_0}{\partial x_i} + \sum_{ij}\xi_{ij}\left(\frac{\partial S}{\partial x_i}\frac{\partial v_0}{\partial x_j} + \frac{\partial S}{\partial x_j}\frac{\partial v_0}{\partial x_i} + \frac{\partial^2 S}{\partial x_i \partial x_j}v_0\right) \\
&+ \sum_{ijk}\xi_{ijk}\left(\frac{\partial S}{\partial x_i}\frac{\partial S}{\partial x_j}\frac{\partial v_0}{\partial x_k} + \frac{\partial S}{\partial x_i}\frac{\partial S}{\partial x_k}\frac{\partial v_0}{\partial x_j} + \frac{\partial S}{\partial x_j}\frac{\partial S}{\partial x_k}\frac{\partial v_0}{\partial x_i}\right. \\
&\left. + \left(\frac{\partial S}{\partial x_i}\frac{\partial^2 S}{\partial x_j \partial x_k} + \frac{\partial S}{\partial x_j}\frac{\partial^2 S}{\partial x_i \partial x_k} + \frac{\partial S}{\partial x_k}\frac{\partial^2 S}{\partial x_i \partial x_j}\right)v_0\right) + \cdots,
\end{aligned}
$$

where the general law of formation is obvious. But on account of the symmetry relations (4), this may be written

$$
\begin{aligned}
&\sum_i \xi_i \frac{\partial v_0}{\partial x_i} + 2\sum_{ij}\xi_{ij}\frac{\partial S}{\partial x_i}\frac{\partial v_0}{\partial x_j} + 3\sum_{ijk}\xi_{ijk}\frac{\partial S}{\partial x_i}\frac{\partial S}{\partial x_j}\frac{\partial v_0}{\partial x_k} + \cdots \\
&\quad + \frac{1}{2}\left(2\cdot 1\sum_{ij}\xi_{ij}\frac{\partial^2 S}{\partial x_i \partial x_j} + 3\cdot 2\sum_{ijk}\frac{\partial S}{\partial x_i}\frac{\partial^2 S}{\partial x_j \partial x_k} + \cdots\right)v_0.
\end{aligned}
$$

But the coefficient of $\partial v_0/\partial x_i$ in the first line is $\partial P/\partial y_i$ if we write $y_i = \partial S/\partial x_i$ in P; and the coefficient of $\partial^2 S/\partial x_i \partial x_j$ in the second line is $(v_0/2)\partial^2 P/\partial y_i \partial y_j$. Hence the required condition for $k=1$ may be written in the form

$$
\sum_i \frac{\partial P}{\partial y_i}\frac{\partial v_0}{\partial x_i} + \frac{1}{2}\left(\sum_{ij}\frac{\partial^2 P}{\partial y_i \partial y_j}\frac{\partial^2 S}{\partial x_i \partial x_j} + Q\right)v_0 = 0. \tag{8}
$$

Let us next throw this linear differential equation in v_0 into

a different form, by use of the curves $x_i = x_i(\tau)$, $(i = 1, \cdots, n)$, defined by the n ordinary differential equations of the first order

$$\frac{dx_i}{d\tau} = \frac{\partial P}{\partial y_i}, \qquad (i = 1, \cdots, n). \tag{9}$$

Along any such curve in n-dimensional $(x_1 \cdots x_n)$-space, (8) may be written

$$\frac{dv_0}{d\tau} + \Phi v_0 = 0, \tag{10}$$

where Φ is the coefficient of v_0 in (8). The solution of (8) is therefore

$$v_0 = v_0^* e^{-\int \Phi d\tau}. \tag{11}$$

Thus the equation for $k = 1$ may be looked upon as determining the value of v_0 throughout a tube of these integral curves, once v_0 has been assigned values v_0^* on a particular transversal surface Σ. The later coefficients do not enter at this first stage $k = 1$.

If now we turn to the later equations for any k $(k > 1)$, it is clear that these have a similar form

$$\frac{dv_{k-1}}{d\tau} + \Phi v_{k-1} + A_{k-1} = 0, \tag{12}$$

where A_{k-1} is a known linear differential expression in v_0, $v_1, \cdots, v_{k-2}$. Hence we find that $v_0, v_1, \cdots$ are determined in succession by their values on the transversal surface Σ.

In particular we may suppose that $v_0, v_1, \cdots$ are given arbitrarily on a small region σ of Σ, continuous together with all of their partial derivatives in σ but vanishing along the boundary of and outside of σ. Evidently these functions will then vanish similarly all along the tube and outside of it. We shall refer to a formal solution ψ of this nature, as an "asymptotic wave packet" solution for obvious reasons.

3. *The Associated Canonical Equations.* Let us assume now that S is not only a solution of the partial differential equation $P = 0$, but belongs to an n-parameter family of solutions

$$S(x_1, \cdots, x_n; c_1, \cdots, c_{n-1}) + c_n,$$

involving n constants $c_1, \cdots, c_n$, one of which, c_n, is additive. We shall assume that the $n-1$ constants $c_1, \cdots, c_{n-1}$ are "independent" in the sense that if we write $y_i = \partial S/\partial x_i$ so that always, by (6), $P(x_1, \cdots, x_n; y_1, \cdots, y_n) = 0$, the y_i's may be regarded as independent except for the relation just written, that is, we shall assume that the $n \times (n-1)$ matrix

$$\left|\left| \partial^2 S/\partial x_i \partial c_j \right|\right|, \qquad (i = 1, \cdots, n; j = 1, \cdots, n-1),$$

is of rank $n-1$. In general an arbitrary (non-singular) solution S of (5) can be imbedded in such a "complete solution."

But by differentiation of the equation $y_i = \partial S/\partial x_i$ along a curve yielding a solution of (9) for any particular set of values of $c_1, \cdots, c_{n-1}$, we obtain

$$\frac{dy_i}{d\tau} = \sum_j \frac{\partial^2 S}{\partial x_i \partial x_j} \frac{\partial x_j}{\partial \tau} = \sum \frac{\partial^2 S}{\partial x_i \partial x_j} \frac{\partial P}{\partial y_j}$$

by use of (9). On the other hand, by partial differentiation of the identity (6) as to x_i, we obtain

$$\frac{\partial P}{\partial x_i} + \sum_{j=1}^{n} \frac{\partial P}{\partial y_i} \frac{\partial^2 S}{\partial x_i \partial x_j} \equiv 0.$$

By combination of this equation and the one which precedes it, we conclude that the following equations also obtain:

$$\frac{dy_i}{d\tau} = -\frac{\partial P}{\partial x_i}, \qquad (i = 1, \cdots, n). \tag{13}$$

It will be observed then that the equations (9), (13) yield a canonical system in x_i, y_i, $(i = 1, \cdots, n)$, of the $2n$th order,

$$\frac{dx_i}{d\tau} = \frac{\partial P}{\partial y_i}, \qquad \frac{dy_i}{d\tau} = -\frac{\partial P}{\partial x_i}, \quad (i = 1, \cdots, n), \tag{14}$$

having a principal function P not containing the independent variable τ.

Moreover we have

$$\frac{d}{d\tau}\left(\frac{\partial S}{\partial c_i}\right) = \sum_j \frac{\partial^2 S}{\partial c_i \partial x_j} \frac{dx_j}{d\tau} = \sum_j \frac{\partial^2 S}{\partial c_i \partial x_j} \frac{\partial P}{\partial y_j} = 0,$$

since $\partial P/\partial c_i \equiv 0$. Hence we infer that the solutions of (14) under consideration satisfy the equations

$$d_i = \frac{\partial S}{\partial c_i}, \qquad (i = 1, \cdots, n-1). \tag{15}$$

But $x_1, \cdots, x_n, y_1, \cdots, y_n$ can evidently be taken as any point on the $(2n-1)$-dimensional manifold $P=0$ at $\tau=0$ so that we obtain in this way the general solution $P=0$ of (14) in the form

$$y_i = \frac{\partial S}{\partial x_i}, \quad (i = 1, \cdots, n); \quad d_i = \frac{\partial S}{\partial c_i}, \quad (i = 1, \cdots, n-1), \tag{16}$$

where $S(x_1, \cdots, x_n, c_1, \cdots, c_{n-1}) + c_n$ is a solution of the multiplier equation of the specified generality. It will be observed that there are $2n-2$ arbitrary constants involved, namely $c_1, \cdots, c_{n-1}, d_1, \cdots, d_{n-1}$. Thus (16) defines a $(2n-2)$-parameter family of curves filling up the $(2n-1)$-dimensional manifold $P=0$; the parameter τ is then determined by setting

$$d\tau = \frac{dx_1}{\partial P/\partial y_1} = \cdots = -\frac{dy_n}{\partial P/\partial x_n},$$

and integrating.

It need hardly be remarked that the equations (14) define the Cauchy characteristics of the partial differential equation $P=0$, in the theory of the solution of which the equations (16) play a well known fundamental part. Thus we may state the following result.

As τ varies, an asymptotic wave packet solution of $L(\psi, \lambda)=0$, belonging to a non-singular phase function S and an amplitude function v_0, travels along the corresponding Cauchy characteristic in $x_1 \cdots x_n$ space, defined by the canonical equations (14), *where the initial values of y_i are given by the equations*

$$y_i = \partial S/\partial x_i, \qquad (i = 1, \cdots, n).$$

4. *On Certain Integral Invariants.* We propose next to give the condition (8) an alternative integral invariant form. It is well known that an integral such as

$$I = \int_{V(\tau)} G v_0^k \, dx_1 \cdots dx_n \tag{17}$$

will be invariant as τ changes in case a certain divergence vanishes:

$$\sum_i \frac{\partial}{\partial x_i}\left(Gv_0^k \frac{\partial P}{\partial y_i}\right) \equiv 0.$$

Here $V(\tau)$ is a volume in $x_1 \cdots x_n$ space which moves as τ changes in accordance with equations (14), where $y_i = \partial S/\partial x_i$. This yields

$$\sum_i \frac{\partial G}{\partial x_i}\frac{\partial P}{\partial y_i}v_0^k + kG\left(\sum_i \frac{\partial P}{\partial y_i}\frac{\partial v_0}{\partial x_i}\right)v_0^{k-1} + G\left(\sum_i \frac{\partial^2 P}{\partial x_i \partial y_i} + \sum_{ij} \frac{\partial^2 P}{\partial y_i \partial y_j}\frac{\partial^2 S}{\partial x_i \partial x_j}\right)v_0^k \equiv 0,$$

which by virtue of (8) reduces to

$$\sum_i \frac{\partial P}{\partial y_i}\frac{\partial G}{\partial x_i} + \left(\sum_i \frac{\partial^2 P}{\partial x_i \partial y_i} + \left(1 - \frac{k}{2}\right)\sum_{ij} \frac{\partial^2 P}{\partial y_i \partial y_j}\frac{\partial^2 S}{\partial x_i \partial x_j} - kQ\right)G \equiv 0. \tag{18}$$

In this equation $x_1, \cdots, x_n$ are the independent variables since $\partial S/\partial x_i$ is substituted for y_i throughout. It may be noticed here that $G = \text{const.}$ is a solution for $k = 2$, $Q = 0$, provided that

$$\sum_i \frac{\partial^2 P}{\partial x_i \partial y_i} \equiv 0, \tag{19}$$

regardless of the choice of y_i; we shall have occasion to employ this result later on.

Conversely it is evident that in general if for the given S and k, a function G is a solution of (18), and if $\int_{V(\tau)} Gv_0^k dx_1 \cdots dx_n$ is an integral invariant for any region V, then v_0 must satisfy the equation (8).

We may now announce the following result.

If v_0 is any amplitude function of the non-singular phase function S and if G satisfies the linear partial differential equation (18) *for this S, and some k, then*

$$I = \int_{V(\tau)} Gv_0^k \, dx_1 \cdots dx_n$$

is an invariant integral for any volume $V(\tau)$, where $x_1, \cdots, x_n$ vary with τ in accordance with (14) *(with $y_i = \partial S/\partial x_i$). Conversely, if this integral is invariant for all regions $V(\tau)$, then v_0 is an amplitude function which satisfies the equation* (8).

5. *The Schrödinger Wave Equation.* Suppose now that we take $n+1$ variables $t, x_1, \cdots, x_n$, with

$$P \equiv \frac{\partial S}{\partial t} + H\left(x_1, \cdots, x_n; \frac{\partial S}{\partial x_1}, \cdots, \frac{\partial S}{\partial x_n}\right),$$

so that the "multiplier equation" takes the form of the usual Hamiltonian equation. The corresponding principal equation $L_0(\psi) = 0$ is then the usual Schrödinger wave equation

$$(20)\quad \frac{2\pi i}{h}\frac{\partial \psi}{\partial t} + H\left(x_1, \cdots, x_n; \frac{2\pi i}{h}\frac{\partial}{\partial x_1}, \cdots, \frac{2\pi i}{h}\frac{\partial}{\partial x_i}\right)\psi = 0,$$

provided we take $\lambda = 2\pi i/h$.

The Schrödinger equation is therefore merely the principal equation which has the usual Hamilton-Jacobi partial differential equation as its multiplier equation, with $\lambda = 2\pi i/h$.

Furthermore, the corresponding Hamiltonian equations are

$$\frac{dt}{d\tau} = 1, \quad \frac{d\left(\frac{\partial S}{\partial t}\right)}{d\tau} = 0,$$

together with

$$(21)\quad \frac{dx_i}{d\tau} = \frac{\partial H}{\partial y_i}, \quad \frac{dy_i}{d\tau} = -\frac{\partial H}{\partial x_i}, \quad (i = 1, \cdots, n),$$

with $y_i = \partial S/\partial x_i$. Hence we find $\tau = t + \text{const.}$, $\partial S/\partial t = \text{const.}$ along any trajectory. But in this case a complete solution can be found in the form

$$S = S^*(x_1, \cdots, x_n; c_1, \cdots, c_{n-1}) + c_n t + c_{n+1}$$

with $y_i = \partial S^*/\partial x_i$ for $i = 1, \cdots, n$. Hence the equations for the Cauchy characteristics reduce to the ordinary Hamiltonian equations(21) with $t = \tau$, associated with Schrödinger's wave equation.

Furthermore, it is easily proved that

$$I \equiv \int_{V(\tau)} \psi\bar{\psi} dt dx_1 \cdots dx_n,$$

where $\bar{\psi}$ denotes the conjugate of ψ, is then an integral invariant to the first order in $1/\lambda$, at least if

$$\sum_i \frac{\partial^2 H}{\partial x_i \partial y_i} \equiv 0. \tag{19'}$$

In fact, since $\lambda = 2\pi i/h$ is a pure imaginary, we have

$$\psi \sim e^{\lambda S} v_0, \quad \bar{\psi} \sim e^{-\lambda S} v_0$$

for S and v_0 real, so that the above integral reduces essentially to $\int_{V(\tau)} v_0^2 dt\, dx_1 \cdots dx_n$, which is of the form treated above with $G \equiv 1$, $k = 2$; furthermore, we have $Q \equiv 0$ in this case. But when rectangular coordinates are employed, we have also

$$H = V(x_1, \cdots, x_n) + \frac{1}{2} \sum k_i y_i^2,$$

where V is the potential energy and $y_1, \cdots, y_n$ are the momental coordinates corresponding to $x_1, \cdots, x_n$, respectively. Hence (19) and (19') obtain, and the integral I is invariant as stated. Finally, since $d\tau = dt$, it follows at once that $\int_{V(\tau)} v_0^2 dx_1 \cdots dx_n$ also remains constant over any region in $x_1 \cdots x_n$ space. Thus we arrive at the following conclusion.

In the special case of the Schrödinger equation in ψ with rectangular coordinates, the asymptotic wave packets follow the corresponding dynamical trajectories, while the squared amplitude integral $\int |\psi|^2 dx_1 \cdots dx_n$ remains constant over any part of the packet, to terms of the order of h.

6. *On Change of Independent Variables.* Suppose now that we make any change of independent variables

$$\bar{x}_i = f_i(x_1, \cdots, x_n), \qquad (i = 1, \cdots, n).$$

Because of the identities

$$\frac{\partial^{[1]}\psi}{\partial x_i} = \sum_j \frac{\partial \bar{x}_j}{\partial x_i} \frac{\partial^{[2]}\psi}{\partial \bar{x}_j},$$

$$\frac{\partial^{[2]}\psi}{\partial x_i \partial x_j} = \sum_{k,l} \frac{\partial \bar{x}_k}{\partial x_i} \frac{\partial \bar{x}_l}{\partial x_j} \frac{\partial^{[2]}\psi}{\partial \bar{x}_k \partial \bar{x}_l} + \frac{1}{\lambda} \sum_k \frac{\partial^2 \bar{x}_k}{\partial x_i \partial x_j} \frac{\partial^{[1]}\psi}{\partial \bar{x}_k} + \cdots,$$

the components $L_0, L_1, \cdots$ in L do *not* remain individually invariant in the equation (1), and in particular the principal part L_0 will not carry over into the new principal part by the ordinary rules. In fact the coefficients transform by the rules valid for the attached Hamilton-Jacobi equation. Hence Schrödinger's wave equation in the form (20) is only maintained (in general) under a linear transformation of the independent variables.

This fact indicates that any coordinate system from which we start is to be regarded as a privileged absolute system of reference for the Schrödinger wave equation, up to an arbitrary linear transformation.

7. *Linear Systems and the Dirac Equations.* Let us now turn briefly to a system of k homogeneous linear partial differential equations in $\psi_1, \cdots, \psi_k$:

$$(22)\quad L_i(\psi_1, \cdots, \psi_k; \lambda) \equiv L_{i0} + \frac{L_{i1}}{\lambda} + \cdots = 0, \quad (i = 1, \cdots, k).$$

Here $x_1, \cdots, x_n$ are the independent variables, and the same operational symbols have been introduced as above. Now each L_{i0} may be written as a sum, $\sum_j L_{ij0}$, where L_{ij0} contains only the terms of L_{i0} which involve ψ_j. Furthermore there is then a corresponding set of polynomials $P_{ij}(x_1, \cdots, x_n; y_1, \cdots, y_n)$ obtained as in the special case $k=1$ treated above.

If we use formal series solutions,

$$\psi_i = e^{\lambda S}\left(v_{i0} + \frac{v_{i1}}{\lambda} + \cdots\right), \quad (i = 1, \cdots, k),$$

and substitute in the k given equations, the leading terms give us the k equations

$$\sum_j P_{ij}\left(x_1, \cdots, x_n; \frac{\partial S}{\partial x_1}, \cdots, \frac{\partial S}{\partial x_n}\right) v_j = 0, \quad (i = 1, \cdots, k).$$

In order that these be consistent, the "multiplier equation" for the system

$$P \equiv |P_{ij}| = 0$$

must be fulfilled. Furthermore if the "phase function" S satisfies this multiplier equation, then the linear equations just written determine the k functions ψ_i up to a proportionality factor.

For the determination of this proportionality factor v and so of v_{i0}, of $v_{i1}, \cdots$, we might proceed as before with a greater degree of algebraic complication of course. It is sufficient for our purposes, however, to observe that here too the phenomenon of wave packets occurs. In fact by elimination we may reduce the given system in various ways to a single linear differential equation in a single unknown function ψ, linear in ψ_i and their partial derivatives. Its multiplier equation is then essentially $P=0$, *with the same polynomicl* P as before, since the phase functions are the same as before. Hence asymptotic series solutions for ψ having the nature of wave packets exist, associated with this particular P, and so there exist also the corresponding solutions $\psi_1, \cdots, \psi_k$. Thus there exist asymptotic wave packets for the system which follow the Cauchy characteristics belonging to P. The arbitrary proportionality factor in v_{i0} on a transversal surface Σ corresponds to the arbitrary v_0 on Σ in the series for ψ.

Thus it is clear that any *system* of "wave equations" is correlated with a multiplier equation and the allied set of characteristics.

Now the well known work of Sommerfeld showed that the program of Schrödinger leads to a successful theory of the fine structure spectral lines, if one takes account of the special theory of relativity in a natural way; but that it fails to account for certain magnetic properties of the atom. Pauli then substituted a system of two wave equations of the first order for the single Schrödinger equation of the second order so as to bring about the indicated modifications; the multiplier equation $P=0$ obtained is again that necessitated by the special theory of relativity. Finally Dirac obtained a system of four equations of the first order, with multiplier equation $P^2=0$.

Without attempting to analyze the formation of the elegant equations of Dirac, it may be pointed out that the retention of the multiplier equation in unaltered form is of itself sufficient to ensure the proper general form of the characteristic numbers and functions (see the second part of this lecture). It becomes then a very puzzling problem to discover whether the equations of Dirac are to be regarded as more than a set of equations built ad hoc. This is an issue upon which I do not feel myself competent to pronounce.

PART II. THE WENTZEL-BRILLOUIN-KRAMERS METHOD AND ASYMPTOTIC SERIES

8. *Formulation of the Problem.*† In order to simplify the form of statement of the problem, we shall write the wave equation to be considered in the form

$$\frac{d^2\psi}{dx^2} + \lambda^2(E - V(x))\psi = 0, \qquad \left(\lambda^2 = \frac{8\pi^2 m}{h^2}\right), \tag{23}$$

and assume that $V(x)$ is real and analytic for all real values of x, and that it possesses a single (absolute) minimum for $x = x_0$, such that $dV/dx = 0$, $d^2V/dx^2 > 0$ at x_0, while $dV/dx \neq 0$ for $x \neq x_0$. Finally we assume that for $|x|$ large, V admits a convergent series expansion in $1/x$ of the form $a + b/x + \cdots$, so that $\lim V(x) = a$ as x becomes infinite. While this is a somewhat idealized form of the case of physical importance, it will be found that the method proposed in justification of the final formula can be extended without essential modification to the cases of physical interest.

We are concerned with the real solutions ψ of (23) which vanish both at $x = -\infty$ and $x = +\infty$, for real and positive λ. Now, for $E \leqq V_0$, the coefficient of ψ in (23) is everywhere negative or 0, and elementary oscillation theorems show that no solution ψ can vanish for $x = \pm\infty$. On the other hand, if $E > a$, this coefficient is everywhere negative, and all solutions oscillate indefinitely often with an amplitude that need not approach 0; this corresponds to the possibility of a continuous spectrum. Hence we may assume that E exceeds V_0 but is less than a.

Let us for the present consider λ as a large positive parameter while E is taken to be restricted as stated. We have then a boundary value problem of classical type, but with the difficulty arising from the singular nature of the boundary conditions (the boundaries lie at infinity) and from the fact that the coefficient of ψ changes sign twice. As far as I know problems of this singular type have not as yet been treated (see, however, Langer, loc. cit.).

† We follow A. Zwaan and J. L. Dunham (loc. cit.) in using a complex variable x. Zwaan's treatment is extremely suggestive, although lacking in essential respects.

9. *An Auxiliary Lemma.* In order to proceed further we shall need the following lemma.

LEMMA. *In any leaf-shaped region σ of the complex x-plane in which $V(x)$ is analytic and which can be covered by a regular family of curves from two points P and Q of its boundary, in such wise that*

$$\Re((V(x) - E)^{1/2}dx) \neq 0 \tag{24}$$

along each curve,† there will exist two solutions $\psi_i(x, \lambda)$, $(i=1, 2)$, analytic in x and λ, and asymptotically represented by the usual formal series solutions $s_i(x, \lambda)$, $(i=1, 2)$, throughout σ.

This lemma is a special case of an obvious extension of results contained in my doctoral thesis.‡

10. *Application of the Lemma.* By use of the above lemma, it is possible to determine such regions σ. We restrict attention to the neighborhood of the axis of reals in the x-plane, and let $x=\alpha$ and $x=\beta$ $(\alpha<x_0<\beta)$ denote the two values of x for which $V(x)-E$ vanishes, so that $V(x)-E$ is positive or negative according as x lies outside of or within (α, β). For $x=\alpha$ this function decreases, while for $x=\beta$ it increases.

Let us make a cut in the complex x-plane from $-\infty$ to α, and from β to $+\infty$, along the axis of reals and consider $(V(x)-E)^{1/2}$ in the cut plane; here we take the positive branch on the upper side of the cut $(-\infty, \alpha)$, and then determine this function throughout the cut plane. Evidently $(V(x)-E)^{1/2}$ is a pure imaginary quantity with negative coefficient of i on the real axis between α and β, and is a negative real quantity on the upper side of the cut $(\beta, +\infty)$. On the lower side of the two cuts, the function is of course equal to the negative of its value at the same point on the upper side.

---

† $\Re$ denotes the "real part of"; E is regarded as fixed.

‡ *On the asymptotic character of the solutions of certain differential equations containing a parameter*, Transactions of this Society, vol. 9 (1908), pp. 219–230. It will be found that for an equation of order n with real parameter ρ (notation of my paper) it is sufficient that along the curves PQ in σ we have

$$\Re(w_1dx) \geqq \Re(w_2dx) \geqq \cdots \geqq \Re(w_ndx).$$

This convenient condition for maintenance of asymptotic form in σ was known to me in 1908, and is indeed obvious from my paper.

Now let us consider the curves

$$\Re((V(x) - E)^{1/2}dx) = 0, \tag{25}$$

which evidently play an important part in the determination of possible regions σ. Under the assumptions made above, the general nature of these curves near the x axis is readily seen to be that indicated in the figure below.

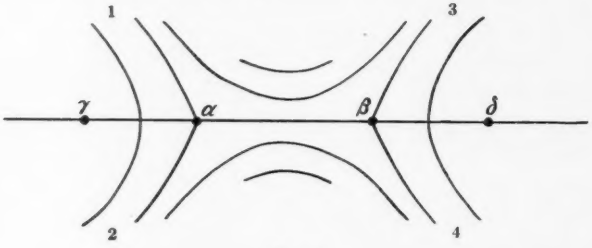


FIG. 1

This leads at once to five special types of regions σ: I: σ above $2\alpha\beta3$; II: σ above $1\alpha\beta4$; III: σ below $1\alpha\beta4$; IV: σ below $2\alpha\beta3$; V: σ between $1\alpha2$ and $3\beta4$.

Here the boundaries of these regions are usually to be excluded. *We shall, however, assume that the point P or Q of the Lemma in regions σ of types* I–IV *may be taken at an infinite end of the real axis.* A critical investigation of the validity of this reasonable assumption is being undertaken by Mr. A. C. Galbraith at Harvard University.

11. *The Distribution of the Characteristic Values.* With these preliminaries in hand we are prepared to determine the distribution of the characteristic values λ. In the first place we are only interested in the solutions ψ of (23) which tend to 0 as x tends to $\pm\infty$. Now we have two formal solutions

$$s_1(x, \lambda) \sim e^{\lambda\int^x (V-E)^{1/2}dx} t_1(x, \lambda), \qquad s_2(x, \lambda) \sim e^{\lambda\int^x (V-E)^{1/2}dx} t_2(x, \lambda),$$

where we take x on the upper side of the cut $(-\infty, \alpha)$ to begin with, and $(V-E)^{1/2}$ as positive there. Here the $t_i(x, \lambda)$ are ordinary power series in $1/\lambda$, and we may suppose that s_1 and t_1 go into s_2 and t_2, respectively, as we traverse the cuts.

Now the unique formal solution of the first type above which reduces formally to 1 for $x=\gamma$ (see the figure) is clearly

$$s_1(x, \lambda)/s_1(\gamma, \lambda) = e^{\lambda \int_\gamma^x (V-E)^{1/2} dx} t_1(x, \lambda)/t_1(\gamma, \lambda).$$

According to the Lemma there is a corresponding ψ_1 having this asymptotic form in the region I, which will evidently approach 0 exponentially as x approaches $-\infty$. Similarly there will exist a solution ψ_2 represented by $s_2(x, \lambda)/s_2(\gamma, \lambda)$ which approaches ∞ under the same circumstances. Hence $\psi_1(x, \lambda)$ is essentially the only solution (up to a constant multiplier) which remains finite as x approaches $-\infty$, and $\psi_1(x, \lambda)/\psi_1(\gamma, \lambda)$ is a special solution, $\psi^*(x, \lambda)$, with the same asymptotic form, which reduces to 1 for $x=\gamma$. This solution is clearly real for x real, since $V(x)$ is real for x real. Hence ψ^* is represented asymptotically by $s_1(x, \lambda)/s_1(\gamma, \lambda)$ in the combined region I+III, since for x below the real axis, ψ^* is conjugate to its value at the conjugate point above the axis, and $s_1(x, \lambda)/s_1(\gamma, \lambda)$ has the same formal property. Taking account of the cut, however, we have

$$\psi_1^*(x, \lambda) \sim \begin{cases} s_1(x, \lambda)/s_1(\gamma, \lambda) \text{ in I, above,} \\ s_2(x, \lambda)/s_2(\gamma, \lambda) \text{ in III, below.} \end{cases}$$

Similarly in the regions II+IV we are led to fix attention upon a solution

$$\psi_2^*(x, \lambda) \sim \begin{cases} s_1(x, \lambda)/s_1(\delta, \lambda) \text{ in II, above,} \\ s_2(x, \lambda)/s_2(\delta, \lambda) \text{ in IV, below,} \end{cases}$$

as yielding the only possible solution which approaches 0 as x becomes positively infinite. It is to be noted that $s_1(x, \lambda)$ is representative of the formal solution which is asymptotic to 0 as x approaches $+\infty$ along the upper side of (β, ∞). Moreover, for a characteristic value and only then, these two solutions must be proportional; that is, their ratio must reduce to essentially the same function $f(\lambda)$ in the overlapping parts above $1\alpha\beta 3$, and below $2\alpha\beta 4$. Hence we are led to the necessary relation

$$\frac{s_1(\delta, \lambda)}{s_1(\gamma, \lambda)} \sim \frac{s_2(\delta, \lambda)}{s_2(\gamma, \lambda)}. \tag{26}$$

Consequently if we write

$$s_i(x, \lambda) = e^{\int^x (d \log \psi_i / dx) dx}, \qquad (i = 1, 2),$$

so that ψ_i is a formal solution of the wave equation (23), the above relation yields

$$\int_\gamma^\delta \frac{d \log \psi_1(x)}{dx} dx + \int_\gamma^\delta \frac{d \log \psi_2}{dx} dx \sim 2k\pi i, \qquad (k \text{ an integer}),$$

or, more briefly, since $d \log \psi_1$ changes to $d \log \psi_2$ as we traverse the cut,

$$\oint \frac{d \log \psi_1}{dx} dx = 2k\pi i,$$

where the path of integration is a positive loop around the points α and β. Written out explicitly this gives the series

$$(27) \quad \oint (V(x) - E)^{1/2} dx + \cdots = \frac{(k + \frac{1}{2})h}{(2m)^{1/2}}, \quad (k = 0, 1, \cdots),$$

which is essentially the desired Wentzel-Brillouin-Kramers equation.

We shall not attempt to consider the degree of precision with which this equation leads to satisfactory approximations to the energy levels $E_0, E_1, \cdots$. It may be anticipated, however, that the approximation will be good whenever the terms of the series on the right diminish rapidly; and that an exact result is obtained in case the series involved converges.

HARVARD UNIVERSITY

RECENT PROGRESS ON WARING'S THEOREM AND ITS GENERALIZATIONS*

BY L. E. DICKSON

1. *Introduction and Summary.* The simplest theorem in question states that every positive integer is a sum of four integral squares. This is an example of a universal Waring theorem. The elaborate theory due to Hardy and Littlewood yields a number $C(s, n)$ beyond which every integer is a sum of s integral nth powers greater than or equal to zero. Since C is excessively large, their theory yields essentially only asymptotic theorems.

For several years the writer has been elaborating his idea that it is possible to supplement these asymptotic theorems and show that they hold also for all integers below C. The resulting new universal theorems are here first published.

The various aspects of Waring's problem may be compared with those of the theory of functions of a complex variable. Such a function may be studied in the neighborhood of infinity (corresponding to asymptotic Waring theorems), or in the neighborhood of the origin (compare our new results in §2), or over the whole plane (corresponding to universal Waring theorems). Analytic continuation of a function has its analog in our extension of a range for which s nth powers suffice to a larger range for which also s nth powers suffice. Such an extension is different from ascent to a still larger range for which $s+1$ nth powers suffice (§3).

2. *On the Ideal Limit for the Universal Waring Theorem.* It has been proved that every positive integer is a sum of nine cubes, and noted that 23 requires nine cubes (8, 8 and seven 1's), whence 9 is the ideal limit for cubes. For biquadrates (or fourth powers), 79, 159, 239, 319, 399, 479 and 559 require 19 biquadrates. The published table to 4100 shows that 19 biquadrates suffice to 4100. But the best theorem to date is that all integers are sums of 35 biquadrates. Hence it is only on evidence by tables that 19 is called the ideal limit for biquadrates.

By a decomposition (into nth powers $\geqq 0$), we mean a linear

* An address delivered at Chicago on June 23, 1933, before the Society and Section A of the American Association for the Advancement of Science.

homogeneous function of 1, $a=2^n$, $b=3^n$, $c=4^n$, $\cdots$, with integral coefficients $\geqq 0$. The sum of the coefficients will be called the weight of the decomposition.

We shall prove results which indicate that the ideal limit for nth powers is $I=q+2^n-2$, where q is the greatest integer $<(3/2)^n$. Thus

$$b = qa + r, \qquad (0 < r < a). \tag{1}$$

Consider $P=q2^n-1$. By (1), $P<3^n$, whence any decomposition of P involves only 1 and 2^n. But P requires more than I such terms except when there are exactly $q-1$ terms 2^n and exactly 2^n-1 terms 1. This proves that P is a sum of I, but not fewer, nth powers.

Besides our results concerning I, which furnish a yard-stick for judging universal theorems, we shall derive various important theorems. In particular, after P there is a rapid reduction below I of the number of nth powers which suffice. This indicates that the known asymptotic results should be much improved in the future.

All integers $>P$ are sums of T_1 nth powers, where T_1 is a certain integer $<I$. Beyond the last one requiring T_1, all are sums of $T_2<T_1$ powers, etc. Let t denote the last $T_i<c$; this t is the largest number (6) which is $<c$.

The following table gives the ideal I, this t, and expresses t as a percent of T. Every integer not exceeding that in the last column is a sum of I nth powers.

n	I	t	t/I	I suffices to
7	143	97	.68	26 244
8	279	152	.54	446 148
9	558	333	.60	4 487 724
10	1079	523	.48	17 006 112
11	2132	1211	.57	167 226 768
12	4223	1523	.36	485 205 633
13	8384	2966	.35	3 124 873 080
14	16673	5245	.31	822 670 468
15	33203	7481	.23	21 293 780 000

Define integers Q and R by the formulas

$$a = (a - r)Q + R, \qquad (0 \leqq R < a - r). \tag{2}$$

Write $\lambda = (Q+R)b$ if $a \geqq q+r+R$, but $\lambda = (a+Q-q-r)b$ if $a < q+r+R$. We shall prove that every positive integer $<\lambda$ is a sum of I integral nth powers $\geqq 0$.

The example $n=8$ will clarify the later theory. Evidently any integer can be expressed as the sum of a number $ma+kb$ and a number chosen from $0, 1, \cdots, 255 = a-1$. These sums are written in the last column of the following tablette:

$$\begin{array}{lll}
66+(m+51)a+(k-2)b & 161+(m+25)a+(k-1)b & \{ma+kb\} \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & & \\
161+ & \{(m+26)a+(k-1)b\} & 95+ \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & & \\
\{(m+52)a+(k-2)b\} & 95+ & 190+ \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & &
\end{array}$$

Here the first (or second) row of dots takes the place of 94 lines obtained by adding $1, 2, \cdots, 94$ in turn to the line above the dots, while for the third row of dots we add $1, 2, \cdots, 65$. Hence there are altogether $95+95+66 = 256 = a$ rows. The three numbers in any line are equal since $b = 25a+161$. The weight of a diagonal number in $\{\ \}$ is less than the weights of the remaining two numbers of the same line. Similar results evidently follow for the rows indicated by dots.

Consider a fourth column obtained from the left one of our three columns by replacing a single b by its value; then a fifth column obtained similarly from this fourth column; etc., until we reach a column free of b. It will be proved that the weight of every number in these added columns exceeds the weight of the number in $\{\ \}$ in the same row and lying in our three columns. In other words, the weight of a number in $\{\ \}$ in the three columns of our tablette is the minimum of the weights of all decompositions of an integer $(<c)$ in the same line. There remains only the simple comparison of these minimum weights for the various values of m and k. The general theory makes this comparison for any n.

The three diagonal elements of our tablette are the values of M_0, M_1, M_2 in the following general theory.

LEMMA. *Let* $M_s=(m+sq+s)a+(k-s)b$, *and* $k\geqq x-1$. *Then* $j+M_0, \cdots, j+M_{x-1}$, $(j=0, \cdots, a-r-1)$, *together give decompositions of all integers* $y+ma+kb$ *for* $y=0, 1, \cdots, x(a-r)-1$.

By (1), M_s is equal to $N_s=s(a-r)+ma+kb$. Also,

$$a-r-1+N_s=N_{s+1}-1.$$

Hence the $j+N_s$ give all integers from N_0 to N_x-1 inclusive.

First, let $k<Q$. The $j+M$ in the lemma with $x=k$, together with $i+M_k$ for $i=0, \cdots, a-1-k(a-r)$, give decompositions of all the integers $z+ma+kb$ for $z=0, \cdots, a-1$. For the maximum i the weight of $i+M_k$ is

$$w_{mk}=a-1-k(a-r)+m+kq+k.$$

For the maximum j, the weight of $j+M_s$, $(0\leqq s\leqq k-1)$, is $a-r-1+m+sq+k$. For the largest value $k-1$ of s, this weight will be $<w_{mk}$ if $k(a-r)<q+r$. Since $k\leqq Q-1$, the latter follows from

$$r-1=(a-r)(Q-1)+R-1, \tag{3}$$

which is derived by subtracting $a-r+1$ from (2). Hence w_{mk} is the largest weight of all our decompositions.

For $m=0, \cdots, q-1$, the largest w_{mk} is

$$C_k=a-2-k(a-r)+q+kq+k. \tag{4}$$

For $m=q$, we shall employ $y+qa+kb$ only for $y=0, \cdots, r-1$, since $y=r$ would yield $(k+1)b$ by (1), which exceeds the integers considered in Theorem 1. Previously we had $y=0, \cdots, a-1>r-1$. We have therefore suppressed one or more of the high values of i in $i+M_k$, and increased m from $q-1$ to q. Hence the new greatest weight is $\leqq C_k$. This proves the following result.

THEOREM 1. *For* $k<Q$ *every integer* $\geqq kb$ *and* $<(k+1)b$ *has a decomposition whose weight is* $\leqq C_k$.

Second, let $k\geqq Q$. By the lemma with $x=Q$ we see that

$$\begin{aligned} &j+M_0, \cdots, j+M_{Q-1}, \quad (j=0, \cdots, a-r-1), \\ &i+M_Q, \qquad\qquad (i=0, \cdots, R-1), \end{aligned} \tag{5}$$

together give decompositions of $w+ma+kb$ for $w=0, \cdots, R-1+Q(a-r)=a-1$, by (2). The greatest weight is the larger of

$$b_{mk} = a - r - 1 + m + (Q-1)q + k,$$
$$d_{mk} = R - 1 + m + Qq + k.$$

For $m=0, \cdots, q-1$, the largest b_{mk} or largest d_{mk} is, respectively,

(6) $B_k = a - r - 2 + Qq + k$, $D_k = R - 2 + q + Qq + k$.

For $m=q$, we shall employ $z+qa+kb$ only for $z=0, \cdots, r-1$, since $z=r$ yields $(k+1)b$, which exceeds the integers considered in Theorem 2. In view of (3) and the Lemma for $x=Q-1$, we now employ

$$\begin{aligned} &j + M_0, \cdots, j + M_{Q-2}, \quad (j = 0, \cdots, a - r - 1), \\ &i + M_{Q-1}, \qquad\qquad (i = 0, \cdots, R - 1), \end{aligned} \tag{7}$$

which together give decompositions of $z+qa+kb$, $(z=0, \cdots, r-1)$. The greatest weight is the larger of

$$a - r - 1 + q + (Q-2)q + k < B_k,$$
$$R - 1 + q + (Q-1)q + k < D_k.$$

Evidently $B_k \geqq D_k$ if and only if*

$$a \geqq q + r + R. \tag{8}$$

THEOREM 2. *For $k \geqq Q$, every integer $\geqq kb$ and $<(k+1)b$ has a decomposition whose weight is $\leqq L_k$, where L_k is the larger of* (6), *whence $L_k = B_k$ if $a \geqq q+r+R$, and $L_k = D_k$ if $a < q+r+R$.*

Employing (3), we see that

$$C_{Q-1} = a - 3 + R - r + Q(q+1). \tag{9}$$

Thus $C_{Q-1} \geqq B_k$ if and only if $k \leqq R+Q-1$. Hence if $a \geqq q+r+R$ and $Q \leqq k \leqq R+Q-1$, Theorem 2 shows that every integer $\geqq kb$ and $<(k+1)b$ has a decomposition whose weight is $\leqq C_{Q-1}$. Apply this result for $k=Q, Q+1, \cdots, R+Q-1$ in turn. We obtain the following theorem.

---

* Except for $n=3$, 10, 11, 12, (8) holds for $2 \leqq n \leqq 27$. Note that $Q=1$ (whence $R=r$) if and only if $a>2r$.

THEOREM 3. *If $a \geqq q+r+R$, every integer $\geqq Qb$ and $<(Q+R)b$ has a decomposition whose weight is $\leqq C_{Q-1}$.*

We shall omit the proof (by induction on n) that, in (1),

$$a > q + r + 1, \qquad (n \geqq 3). \tag{10}$$

Hence C_k decreases when k increases, since

$$C_{k-1} - C_k = a - q - r - 1 \geqq 1. \tag{11}$$

We have $C_0 = a+q-2 = I$, $C_1 = r+2q-1$, $C_2 = 2r-a+3q$.

If an integer h is a sum of C nth powers $\geqq 0$, we shall say that C suffices for h. Evidently G suffices for h if $G > C$.

THEOREM 4. *For $k = 0, \cdots, Q-1$, C_k suffices for all integers $\geqq kb$ and $<Qb$.*

This is true by Theorem 1 if $k = Q-1$. To proceed by induction from k to $k-1$, let Theorem 4 hold for a certain k. Then C_{k-1} (which exceeds C_k) suffices for integers $\geqq kb$ and $<Qb$. By Theorem 1, C_{k-1} suffices for integers $\geqq (k-1)b$ and $<kb$. The two results show that C_{k-1} suffices for integers $\geqq (k-1)b$ and $<Qb$, whence Theorem 4 holds when k is replaced by $k-1$.

Theorems 3 and 4 yield the following theorem.

THEOREM 5. *Let $a \geqq q+r+R$. For $k = 0, \cdots, Q-1$, C_k suffices for all integers $\geqq kb$ and $<(Q+R)b$. In particular, $C_0 = I$ suffices for all integers $<(Q+R)b$. If $k \geqq Q$, B_k suffices for all integers $\geqq Qb$ and $<(k+1)b$.*

The final statement follows from Theorem 2.

Finally, let $a < q+r+R$. In Theorem 2, $L_k = D_k$. We shall have $C_{Q-1} \geqq D_k$ if $k \leqq a-1+Q-q-r$. Hence, as in the proof of Theorem 3, every integer $\geqq Qb$ and $<(a+Q-q-r)b$ has a decomposition whose weight is $\leqq C_{Q-1}$. Combining this with Theorem 4, we have the following additional theorem.

THEOREM 6. *Let $a < q+r+R$. For $k = 0, \cdots, Q-1$, C_k suffices for all integers $\geqq kb$ and $<(a+Q-q-r)b$. If $k \geqq Q$, D_k suffices for all integers $\geqq Qb$ and $<(k+1)b$.*

We shall define a tablette T with $Q+1$ columns whose $a-1$ diagonal elements are (5) and which reduces to our special tablette when $n = 8$. Let $s \leqq Q$. Write D for $(s-t)(a-r)$. To the right of M_s occur the elements $D + M_t$, $(t = 0, \cdots, s-1)$, all of

which are equal to M_s by (1). Since $1 \leqq s-t \leqq Q$, it follows that $D>0$ and, by (2), $D \leqq a-R<a$. Hence the weight of $D+M_t$ is $D+m+tq+k$, which will exceed the weight $m+sq+k$ of M_s if $D>(s-t)q$, and hence if $a-r>q$, which is true by (10).

To the left of M_t occur the elements

$$a - D + (m + sq + s - 1)a + (k - s)b \ (= M_s - D)$$

for $s=t+1, \cdots, Q$. We saw that $D \leqq a-R$, $D>0$. Hence $R \leqq a-D<a$. The weight of the displayed number is $a-D+m+sq-1+k$. This will exceed the weight $m+tq+k$ of M_t if $a-D-1+(s-t)q>0$. But its first part is $\geqq 0$ and second part is >0.

THEOREM 7. *In every row of tablette T the least weight is that of the diagonal element M_s.*

We annex columns to the left of T exactly as for $n=8$. Write

$$F_{hi} = (m + hQq + hQ + iq + i - h)a + (k - hQ - i)b.$$

Since $(Qq+Q+q)a=a-r-R+(Q+1)b$ follows by eliminating b by (1) and applying (2), we see by cancellation that

$$F_{hi} = a - r - R + F_{(h-1),(i-1)}. \tag{12}$$

The weight of F_{hi} is $m+k+hQq+iq-h$. This exceeds the weight of the second member of (12) if and only if

$$Qq + q - 1 + R > a - r. \tag{13}$$

Hence if (13) holds, F_{hi} is not a minimum decomposition when $h>0$, $i>0$. Also $F_{0i}=M_i$, which is in T for $i=0, \cdots, Q$. Hence for new columns we have $h \geqq 1$, $i \geqq 1$. Since F_{hi} is then not a minimum decomposition, the same is evidently true of all $j+F_{hi}$. The number at the top of the column containing F_{hi} is the sum of an integer $\geqq 1$ and the function obtained from F_{hi} by subtracting unity from the coefficient of a. Hence the entries above F_{hi} are not minimum decompositions.

THEOREM 8. *In case* (13), *all minimum decompositions of numbers $<c$ occur in T.*

Condition (13) holds if $n=2$, 3, 5, 6, 8–12, 14, 15, but for no further $n<36$. For a value of n, other than these eleven, it may happen that some number smaller than B_k suffices, etc., so that

our theory is only a first approximation. But a more refined and much more elaborate theory would add no new information about the ideal I.

The numbers (6) were obtained as the weights of $j+M_s$ for $j=a-r-1$, $s=Q-1$, and of $i+M_Q$ for $i=R-1$, each with $m=q-1$, namely,

(14) $a-r-1+(Qq+Q-2)a+(k-Q+1)b$, of weight B_k;

(15) $R-1+(q-1+Qq+Q)a+(k-Q)b$, of weight D_k.

The proof leading to Theorem 2 evidently shows also that if $a>q+r+R$, then B_k-1 suffices from kb to (14), exclusive, while if $a<q+r+R$, then D_k-1 suffices from kb to (15), exclusive. Evidently $B_{k-1}=B_k-1$.

Since the expectation is that a smaller number suffices after c than before c, we shall choose k to be the largest integer for which (14), or (15), is $<c$. In the preceding paragraph we replace k by $k+1$ and conclude that B_k or D_k suffices to c.

The number (4) was obtained as the weight of $i+M_k$ for $m=q-1$:

(16) $a-1-k(a-r)+(q-1+kq+k)a$, of weight C_k,

$(k=0,\cdots,Q-1)$.

n	a	q	r	$a-r$	Q	R
5	32	7	19	13	2	6
6	64	11	25	39	1	25
8	256	25	161	95	2	66
9	512	38	227	285	1	227
10	1 024	57	681	343	2	338
11	2 048	86	1 019	1029	1	1019
12	4 096	129	3 057	1039	3	979
14	16 384	291	15 225	1159	14	158
15	32 768	437	29 291	3477	9	1475

Case $n=8$. Thus $a>q+r+R=252$; (14) is $94+50a+(k-1)b$. Since $c=248+50a+8b$, the largest k is 9. Theorem 5 states that $C_0=279$ suffices for all positive integers $<68b=446148>5^8$, $C_1=210$ suffices for all $\geqq b$ and $<68b$, and $B_9=152$ suffices for all $\geqq 2b$ and $<c$. That no fewer suffice to c follows from Theorems

7 and 8. Expressed otherwise, only the following peaks occur from 0 to c:

peak 279 at $255+24a$, peak 210 at $160+50a$, peak 152 at $94+50a+8b$.

Case $n=5$. Peaks $C_0=37$ at $31+6a=223$, $C_1=32$ at $18+14a=466$, $B_3=28$ at $12+14a+2b=946$, $D_3=28$ at $5+22a+b=952$.

Case $n=6$. Peaks $C_0=73$ at $63+10a$, $B_4=52$ at $38+10a+4b$.

Case $n=9$. Peaks $C_0=548$ at $511+37a$, $B_{12}=333$ at $284+37a+12b$.

Case $n=10$. Peaks $C_0=1079$, $C_1=794$, $D_{16}=523$.

Case $n=11$. Peaks $C_0=2132$ at $2047+85a$, $D_{22}=1211$ at $1018+172a+21b$.

Case $n=12$. Peaks $C_0=4223$, $C_1=3314$, $C_2=2405$, $D_{30}=1523$.

Case $n=14$. Peaks $C_0=16673, \cdots, C_k=C_{k-1}-867, \cdots, C_{13}=5402$, $B_{14}=5245$.

Case $n=15$. Peaks $C_0=33203, \cdots, C_k=C_{k-1}-3039, \cdots, C_8=8891$, $B_{73}=7481$.

3. *Formulas for Ascent.* THEOREM 9. *Let a polynomial* $f(x)$ *take integral values* $\geqq 0$ *for every integer* $x\geqq 0$ *and let*

$$f(x+1)-f(x) \tag{1}$$

increase with x. *Make the hypothesis* (H) *that every integer* i *for which* $l<i\leqq g+f(0)$ *is a sum of* $k-1$ *values of* $f(x)$ *for integers* $x\geqq 0$. *Let* m *be an integer (preferably the maximum one) such that*

$$f(m+1)-f(m)<g-l. \tag{2}$$

Then every integer I *for which* $l+f(0)<I\leqq g+f(m+1)$ *is a sum of* k *values of* $f(x)$ *for integers* $x\geqq 0$.

By (1) and (2), we get

$$f(j+1)-f(j)<g-l, \quad (j=0, 1, \cdots, m). \tag{3}$$

For a fixed j consider an integer I for which

$$g+f(j)<I\leqq g+f(j+1). \tag{4}$$

Then $g\geqq I-f(j+1)>g+f(j)-f(j+1)>l$ by (4) and (3). Hence, by (H), $i=I-f(j+1)$ is a sum of $k-1$ values of $f(x)$, whence I is a sum of k values. Apply the latter result for $j=0, \cdots, m$ in turn, and note that each of the $m+1$ intervals (4) ends just where the next begins. Hence every integer which exceeds

$g+f(0)$ and is $\leqq g+f(m+1)$ is a sum of k values. By (H), those from l to g are sums of $k-1$ values. Employ the further value $f(0)$. Hence all from $l+f(0)$ to $g+f(0)$ are sums of k values.

The case $f(x)=x^n$ of Theorem 9 may be stated as follows.

THEOREM 10. *If every integer $>l$ and $\leqq g$ is a sum of $k-1$ integral nth powers $\geqq 0$, and if m is an integer (preferably the maximum one) for which**

$$(m+1)^n - m^n < g - l, \tag{5}$$

then every integer $>l$ and $\leqq g+(m+1)^n$ is a sum of k integral nth powers $\geqq 0$.

There are several results† analogous to the following lemma.

LEMMA. *If n is a positive integer, $l\geqq 0$, $s\geqq l+1$, there exists a positive integer i such that*

$$l \leqq s - i^n < l + ns^{(n-1)/n}. \tag{6}$$

This implies the following result.

THEOREM 11. *Let $L>0$ and let l be an integer $\geqq 0$. Write $\sigma=(L/n)^{n/(n-1)}$. If $\sigma\geqq l$ and if all integers between l and $l+L$, inclusive, are sums of k integral nth powers $\geqq 0$, then every integer s between l and σ, inclusive, is a sum of $k+1$ integral nth powers $\geqq 0$.*

By hypothesis, l is a sum of k and hence of $k+1$ integral nth powers. Hence let $s\geqq l+1$. Thus (6) holds for a positive integer i. For an s in Theorem 11, $s\leqq\sigma$. Hence the final number in (6) is

$$\leqq l + n\sigma^{(n-1)/n} = l + L.$$

By the hypothesis in Theorem 11, $s-i^n$ is therefore a sum of k nth powers, whence s is a sum of $k+1$.

Write $L_0=l+L$, $\nu=(1-l/L_0)/n$. Then Theorem 11 with $\sigma=L_1$ is equivalent to the following statement.

Let l be an integer $\geqq 0$. Let $L_0>l$, $L_1=(\nu L_0)^{n/(n-1)}$. If $L_1\geqq L_0$ and if all integers between l and L_0 inclusive are sums of k nth powers, then all integers between l and L_1 inclusive are sums of $k+1$ nth powers.

* Since $x^n-(x-1)^n<nx^{n-1}$, (5) holds if $m+1$ is $(g-l)/n$ raised to the power $1/(n-1)$.

† Dickson, American Journal of Mathematics, vol. 49 (1927), p. 242; Lemma 2, $t=1$.

Then all between l and $L_2=(\nu L_1)^{n/(n-1)}$ are sums of $k+2$ integral nth powers since

$$L_2/L_1 = (L_1/L_0)^{n/(n-1)} \geqq 1,\ L_2 \geqq L_1,$$

and likewise all between l and

$$L_{t+1} = (\nu L_t)^{n/(n-1)} \tag{7}$$

are sums of $k+t+1$ powers. The following formula

$$\log L_t = \left(\frac{n}{n-1}\right)^t (\log L_0 + n \log \nu) - n \log \nu \tag{8}$$

follows readily from (7) by induction on t. Evidently (8) holds if $t=0$.

THEOREM 12. *Let l be an integer $\geqq 0$. Let*

$$\nu = (1 - l/L_0)/n, \quad L_0 > l, \quad (\nu L_0)^{n/(n-1)} \geqq L_0.$$

Compute L_t by (8). *If all integers between l and L_0 inclusive are sums of k integral nth powers $\geqq 0$, then all integers between l and L_t inclusive are sums of $k+t$ integral nth powers $\geqq 0$.*

4. *New Universal Theorems for Fifth Powers.* A table giving a minimum decomposition of each integer $\leqq 300{,}000$ into fifth powers is in press by the British Association for the Advancement of Science. By the table and deductions from it, we find that all numbers from 98604 (the last peak 17) to 191,263 are sums of 16 fifth powers; all from 191,263 (the last peak 16) to 470,348 are sums of 15; all from 470,348 (the last peak 15) to 786,159 (the last peak 14) are sums of 14; all on to 839,000 are sums of 13.

We shall make no use of the preceding facts, but rely on the following new tables. A table was made from $A=839{,}000$ to 929,000 and used to prove that all numbers from A to $B=1{,}466{,}800$ are sums of 13 fifth powers. Another table from B to 1,600,000 showed that 12 powers suffice except for seven numbers which seem to require 13. It was used to prove that 13 powers suffice from B to $C=3{,}470{,}000$. A final table from C to $D=3{,}600{,}000$ showed that 11 powers suffice except for 13 numbers which seem to require 12. It was used to prove that 12 powers suffice from C to $D+18^5=5{,}489{,}568$. By three applica-

tions of Theorem 10 with $m=24, 40, 70$, we find that 15 powers suffice from A to $L_0=1{,}934{,}501{,}745$. We apply Theorem 12 with $t=39$ and find that all positive integers $\leqq N$ are sums of $15+39=54$ fifth powers if $\log N=34861$. But James has proved that every integer $>C$ is a sum of 54 if $\log C=34251$.

THEOREM 13.* *Every positive integer is a sum of* 54 *fifth powers.* The best earlier result was 58.

5. *New Universal Theorems for nth Powers,* $n=6-12$. Gelbcke applied the method of Winogradow to simplify the proof by Hardy and Littlewood of their best result.† James has determined a constant C for which all the inequalities in that paper hold. Every integer $>C$ is a sum of s nth powers if $\log_e C=20n^3 2^n$. To determine η, let d be the largest integer $\leqq \log_e (n-1)/\log 2$. Define ζ_k by (2) of §7. Write

$$D=(d+2)(n-1)-2^{d+1}+0.1,\quad H=(n-2)2^{n-2}+n,$$

$$B=4+\zeta_k,\quad R=\left\{1+\left(\frac{n-1}{n}\right)^{B-2}\right\}n2^{n-2}-D(B+2).$$

Then

$$\eta=\frac{Ds+R}{s-H-B-2}.$$

For $n=6, 7, 8, 12$, η is

$$\frac{12.1s-114.3}{s-88},\quad \frac{16.1s-186.9}{s-193},$$

$$\frac{20.1s-162}{s-426},\quad \frac{39.1s+9055.4}{s-10335}.$$

R. C. Shook has lately completed extensive tables of minimum decompositions into sixth powers. By them he proved that 23 sixth powers suffice from 2,120,044 to 2,426,148. By Theorems 10 and 12, he found that all integers $\leqq N$ are sums of 160 sixth powers if $\log\log N=11.012$. But $\log\log C=10.890$.

THEOREM 14. *Every positive integer is a sum of* 160 *sixth powers.*‡

* The earlier discussion in this Bulletin, vol. 37 (1931) p. 551, is incomplete since it used the erroneous $\log C=446$.

† Mathematische Annalen, vol. 105 (1931), pp. 637–652.

‡ James had obtained 183. The earlier limit was 478.

For $n=7$, the first five peaks are 143 at $127+16a$, 133 at $116+16a+b$, 123 at $105+16a+2b$, 113 at $94+16a+3b=8703$, and 105 at $83+16a+6b=15253$. In the short stretch on to $c=4^7=16384$ occur two peaks 97. From 16300 to 22038 the largest minimum, 75, is that of $10+59a+6b$. Using only the last result, we find by Theorem 10 that 90 seventh powers suffice to 18,499,724. Then by Theorem 11, 320 suffice to n, where $\log\log n=15.539$. But $\log\log C=15.243$.

THEOREM 15. *Every integer is a sum of* 320 *seventh powers.*

I have reduced 320 to 259. The limit by algebra was 3806.

For $n=8$, we found the three peaks 279, 210, 152 to c. The next peaks are 149 at $65+76a+7b+c$, 137 at $52+76a+7b+2c$, 122 at $52+50a+18b+2c$, and 119 at $l=393{,}095$. We verified that 111 eighth powers suffice from l to $g=1{,}683{,}200$. Thus $6^8-5^8<g-l$. Hence in Theorem 10 the maximum m is 5, whence 112 powers suffice from l to $g+6^8=3{,}362{,}816$. By repeated use of Theorem 10, it was found that 129 eighth powers suffice from $3c=196{,}608$ to $2{,}235{,}617\times10^9$. By Theorem 12 with $t=446$, we find that $129+446=575$ powers suffice to L, where $\log\log L=26.777$. But $\log\log C=26.677$.

THEOREM 16.* *Every positive integer is a sum of* 575 *eighth powers.*

For $n=12$, 1469 powers suffice from $129a+30b=16{,}471{,}614$ to $b+c=17{,}308{,}657$. By Theorem 10, we find that 1560 powers suffice from $3b$ to $5{,}497{,}325\times10^{21}$. By Theorem 12 with $t=9151$, we find that 10711 powers suffice to L if $\log\log L=346.995$. But $\log\log C=346.722$.

THEOREM 17. *Every integer is a sum of* 10711 *twelfth powers.*

Kempner† expressed $t(x^2+y^2+z^2)^6$ as a sum of r 12th powers, where $t=12!/16$. I find that $r=13{,}013{,}280$ and conclude that every integer is a sum of $k=4223+478r$ twelfth powers, where $k=6\frac{1}{4}$ billion. There is no other result by algebra. Starting with very small intervals, James proved this theorem.

* Using merely the fact that 120 suffice from $A=65449$ to $A+a$, inclusive, James verified that every integer is a sum of 595 eighth powers. The earlier limit was 31,353.

† Mathematische Annalen, vol. 72 (1912), p. 398. See [18] below.

THEOREM 18. *Every integer is a sum of* 1177 *ninth powers, and of* 2421 *tenth powers.*

I have reduced 1177 to 981. The earlier result for tenth powers was 140004.

6. *Cubes and Fourth Powers.* Our method has succeeded for $n \geqq 5$, but fails for $n = 3$. The asymptotic theory yields a constant C of about a million digits beyond which every integer is a sum of nine cubes. All integers between 8043 and 40,000 (the limit of tables) are sums of six cubes. Then by Theorem 10 all integers from 8043 to 1,257,375 are sums of seven cubes. By two more ascents, we conclude that all integers $\leqq 2\frac{1}{4}$ billion are sums of nine cubes. While this limit is far below C, it yields the value beyond which the algebraic proof shows that all integers are sums of nine cubes.

Consider fourth powers (biquadrates). If x is odd, x^2 is of the form $8N+1$, whence x^4 is of the form $16m+1$. Thus every biquadrate is $\equiv 0$ or 1 (mod 16), so that no integer $\equiv 15$ (mod 16) is a sum of 14 or fewer biquadrates. Hence every integer $15+16m$ requires at least 15 biquadrates.

We can go further and prove that there are infinitely many integers which require at least 16 biquadrates. If a positive integer K is not a sum of fewer than m biquadrates, where $m \leqq 16$, the same is true of $16K$ and hence of 16^tK, if $t \geqq 1$. For, if $16K$ were a sum of fewer than m and hence fewer than 16 biquadrates, they must all be even, since any biquadrate is $\equiv 0$ or 1 (mod 16). Removing the factor 16, we obtain a contradiction to the hypothesis concerning K. Evidently if K is a sum of m biquadrates x_i^4, then 16^tK is the sum of the $(2^tx_i)^4$. Hence if K has the minimum m, where $m \leqq 16$, then 16^tK has the minimum m. Since $31 = 2^4+15$ has the minimum 16, $16^t \cdot 31$ has the minimum 16 for every $t \geqq 1$.

Bretschneider's table gave a minimum decomposition into biquadrates of each integer <4100. Emily Chandler has recently extended the table to 28,561 and proved that all integers between 13,800 and 158,800 are sums of 16 biquadrates. She concludes that all integers $<10^{26}$ are sums of 19 biquadrates. Then by Theorem 12 with $t = 16$, all integers $\leqq L$ are sums of 35 biquadrates if $\log\log L = 3.37$. By the asymptotic theory all integers $> C$ are sums of 35 if $\log\log C = 3.30$.

THEOREM 19. *All positive integers are sums of* 35 *fourth powers.*

The best result by algebra is 37.

7. *Survey of the Literature on Waring's Problem on Powers.* That every positive integer is a sum of four integral squares was noted by Bachet in 1621, and a proof by descent claimed by Fermat in 1659, while the first published proof was due to Lagrange in 1770. The same year, Edward Waring [1] conjectured that every (positive) integer is a sum of at most 9 (positive integral) cubes, also is a sum of at most 19 biquadrates, and so on.

Except for squares, the first case of a Waring theorem actually proved is due to Liouville [2]. Every positive integer P is a sum of 53 fourth powers. He used an identity equivalent to, but more complicated than, the identity

$$
\begin{aligned}
6(a^2+b^2+c^2+d^2)^2 = {} & (a+b)^4+(a-b)^4+(a+c)^4 \\
& +(a-c)^4+(a+d)^4+(a-d)^4 \\
& +(b+c)^4+(b-c)^4+(b+d)^4 \\
& +(b-d)^4+(c+d)^4+(c-d)^4.
\end{aligned} \tag{1}
$$

Every P is of the form $6q+r$, where $q \geqq 0$, $0 \leqq r \leqq 5$, whence r is a sum of 5 fourth powers each 0 or 1. Also, q is a sum of four squares m_i^2. For each i, $m_i = a^2+b^2+c^2+d^2$, and (1) shows that $6m_i^2$ is a sum of 12 fourth powers. Hence $6q$ is a sum of 4×12, whence P is a sum of $48+5$ fourth powers.

This limit 53 was reduced to 41 by Lucas in 1878, to 39 by Fleck in 1906, to 38 by Landau in 1907, and finally to 37 by Wieferich in 1909.

In 1895, Maillet [3] proved that every positive integer P is a sum of 21 cubes $\geqq 0$, at least 5 of which are 0 or 1, while all sufficiently large P are sums of 17.

In 1909, Hilbert [4] proved that every positive integer P is a sum of N_m integral mth powers $\geqq 0$, where N_m is a finite number (not determined) which depends upon m but not on P. By use of a multiple integral, he first proved the existence (conjectured by Hurwitz [5]) of an identity which expresses $(x_1^2+\cdots+x_5^2)^m$ as a sum of $2m$th powers of linear functions

of $x_1, \cdots, x_s$ with integral coefficients, each power being multiplied by a positive rational number.

In addition to (1), such an identity had been given by Fleck [6] to obtain $N_6 \leqq 184N_3 + 59$; by Hurwitz [5] to get $N_8 \leqq 36{,}119$; by Schur [7] to show that N_{10} is finite [140004].

Wieferich [8] proved that every integer is a sum of nine cubes.

By use of the analytic theory of primes, Landau [9] proved that all sufficiently large integers are sums of eight cubes $\geqq 0$.

Wieferich [10] proved that $N_5 \leqq 59$, $N_7 \leqq 3806$.

Hilbert's proof has been materially simplified by Hausdorff [11] and Stridsberg [12]. An exposition of the resulting proof is due to Oppenheim [13]. By altering Stridsberg's proof at the point where he had used integrals, Remak [14], Frobenius [15], and Stridsberg [16] reduced the proof of Waring's theorem to algebraic processes.

There exist infinitely many positive integers which are not sums of n or fewer positive nth powers (Hurwitz [5] and Maillet [17]). This follows from $S=1$, where S is the superior limit for $k=\infty$ of C/k, C being the number of positive integers $\leqq k$ which are sums of n or fewer positive nth powers.

Kempner [18] proved that $S<1/n!$ and concluded that there is an infinitude of positive integers of each of the forms $9l$, $9l+1, \cdots, 9l+8$, such that each is not a sum of fewer than four positive cubes. Also there is an infinitude of positive integers each not a sum of fewer than nine sixth powers, and an infinitude each not a sum of fewer than 2^{q+2} powers with the exponent 2^q, for $q>1$.

Kempner used a new identity to reduce Fleck's [6] limit for N_6 to 970. While Hurwitz [5] expressed the numbers <5040 as sums of units, Kempner noted that they are sums of at most 273 eighth powers (of 1 or 2) and so reduced the limit for N_8 by $5039-273$. He gave identities which yield limits for N_{12} and N_{14} in terms of N_6 and N_7, respectively. (See below Theorem 17.)

Baer [19] reduced the limit for N_5 to 58 and that for N_6 to 478. He proved that all large numbers of various linear forms are sums of seven [20] (and others of 8) positive cubes.

Hardy and Ramanujan [21] gave an asymptotic formula (involving the zeta and gamma functions) for the number of

ways n is a sum of rth powers of positive integers (rearrangements of the same powers not being counted as distinct).

Of prime importance is the theory originated by Hardy and Littlewood, which applies not merely to Waring's problem but also to various other problems in additive number theory, such as the theory of partitions and Goldbach's and related theorems on sums of primes. The general outlines of their method were explained in their initial paper [22] of 1919. Details of proofs are found in their series of memoirs [23] entitled *Some Problems of Partitio Numerorum*, I–VIII. They obtained asymptotic formulas for the number $r(n, k, s)$ of ways to express a positive integer n as a sum of s kth powers of integers $\geqq 0$. Let $G(k)$ denote the least s such that every sufficiently large integer is a sum of s kth powers. They proved that $G(k) \leqq (k-2)2^{k-1}+4$, whence it suffices to use 9 cubes, 21 biquadrates, 53 fifth powers, etc. In VI, they proved that, if $k \geqq 4$, $G(k) \leqq (k/2-1)2^{k-1} + k + 5 + \zeta_k$, where

$$\zeta_k = \left[\frac{(k-2)\log 2 - \log k + \log(k-2)}{\log k - \log(k-1)}\right]; \tag{2}$$

here $[y]$ denotes the greatest integer $\leqq y$. Hence all large numbers are sums of 19 biquadrates, 41 fifth powers, 87 sixth powers, 193 seventh powers, 425 eighth powers, 949 ninth powers, 2113 tenth powers, etc.

Material simplifications in the proofs have been made by Landau [24], Weyl [25], and Gelbcke [26], who applied the method of Winogradow to prove the results in VI.

In II, the question is raised whether or not the asymptotic result for $r(n, 4, 21)$ holds true also for $r(n, 4, s)$, $s > 21$. This was answered affirmatively by Ostrowski [27], also for $r(n, k, s)$.

Hardy [28] discussed the number of representations as a sum of n cubes.

Western [29] made extensive calculations to find the least G such that all sufficiently large integers are sums of G integral cubes $\geqq 0$. The work of Hardy and Littlewood makes it highly plausible that G is 4 or 5. An examination of 254,000 numbers made the following inferences appear probable.

There are only a finite number of integers not a sum of five cubes, and the greatest of them is $< 2 \cdot 10^6$ and is $\equiv \pm 4 \pmod 9$.

G is 4, and the greatest integer not a sum of 4 cubes lies between 10^{12} and 10^{14}.

Landau [30] wrote $N_k(x)$ for the number of integers $n \leqq x$ which are sums of two kth powers $\geqq 0$. If k is even, the inferior limit for $x=\infty$ of $x^{-2a}N_k(x)$ log x is positive. This implies a similar result on the number of integers $n \leqq x$ which are sums of s kth powers.

Chowla [31] proved that every large integer N is the sum of the cubes of eight integers each $>(1/9)N^{1/3}$; every large N is a sum of eight cubes in at least $gN^{2/9}$ ways, where g is a positive constant.

G. K. Stanley [32] proved that every large integer is a sum of two squares and four cubes $\geqq 0$, also is a sum of one square and six cubes $\geqq 0$.

R. D. James [33] has recently improved considerably the best earlier result for the case of odd powers. All sufficiently large integers are sums of 35 fifth powers (in place of the earlier 41). If k is odd and $\geqq 7$, all sufficiently large integers are sums of

$$(k-3)2^{k-2}+k+9+\zeta_k$$

kth powers, where ζ_k is given by (2). This gives 164 for $k=7$ (instead of the earlier 193), and 824 for $k=9$ (in place of 949).

8. *Sums of Powers with Coefficients.* The writer [34] proved that every integer $\geqq 0$ is a sum of 17 biquadrates and the doubles of ten biquadrates. Also is a sum of r sixth powers and the products of s sixth powers by 8, where either $r=115$, $s=108$, or $r=178$, $s=99$. He proves other similar theorems. Elsewhere he [35] wrote C_n for a sum of n cubes and proved that C_8+tx^3 is universal (represents all integers $\geqq 0$) if $1 \leqq t \leqq 23$, $t \neq 20$, but not if $t>23$. Also $C_7+2x^3+ty^3$ is universal if $1 \leqq t \leqq 34$, except when $t=10, 15, 20, 25, 30$. Also, $C_7+3x^3+ty^3$ is universal if $1 \leqq t \leqq 9$, $t \neq 5$. If N is a positive integer prime to 6, every integer $\geqq 23^3N$ is represented by $6(x^2+y^2+z^2)+Nw^3$. For $l=1, 2, 3, 4$, or 5, C_7+lx^3 represents all sufficiently large integers.

K. C. Yang [36] obtained limits beyond which every integer is represented by C_7+rx^3 for $r=3, 5, 7, 8$. Every odd integer $\geqq k=23 \cdot 10^{14}$ is represented by C_7+rx^3 for $r=2$, 4, or 6. Every integer $\geqq k$ which is double an odd is represented by C_7+2x^3.

R. E. Huston [37] employed $s=(k-2)2^{k-1}+5$, as in the

first Hardy-Littlewood theory, $f=a_1x_1^k+\cdots+a_sx_s^k$, and proved that all sufficiently large integers are represented by f with each x_i an integer $\geqq 0$, if for every integer n there exists a solution of $f\equiv n \pmod{p^\gamma}$ when p is a prime not dividing both a_i and x_i for $i=1, \cdots, s$, and $\gamma=\theta+1$ if $p>2$, $\gamma=\theta+2$ if $p=2$, where p^θ is the highest power of p dividing k. In various general cases, the congruence was proved to have such a solution. This is true for example if $k=3$, $a_1=a_2=1$, $a_3=2, 3, 4$ or 5, with $a_4, \cdots, a_9$ arbitrary; also for $k=4$, if five coefficients are 1, 1, 2, 4, 8, and the remaining 16 are arbitrary.

9. *Remarkable Empirical Generalizations of Waring's Theorem.* Let $(a_1, \cdots, a_r)_n$ denote the form $a_1x_1^n+\cdots+a_rx_r^n$, in which the a's are all positive integers, and call r its *order*, and $a_1+\cdots+a_r$ its *weight*. In §2, we discussed the ideal limit I for a universal theorem on nth powers. Since ax^n is a sum of a nth powers, I is the minimum weight of a form which represents all positive integers. We shall here restrict our attention to forms of weight I.

If a form of weight I represents all positive integers, its order exceeds 4 if $n=3$, 6 if $n=4$, 7 if $n=5$, 9 if $n=6$, and 10 if $n=7$. But there exist forms of one higher order which represent all integers up to a high limit.

Let f be a form which represents k, and let $a_1=s+t$. The form $g=(s, t, a_2, \cdots, a_r)_n$ shall be said to be derived from f by the *partition* of a_1 into $s+t$. If we give to the first two variables in g the same value x_1 that was employed in $f=k$, we see that also g represents k. Hence any form derived from f by partition represents every integer which can be represented by f (and usually represents further integers).

It may be shown [38] in a few lines that, if $n=3$, the only forms of weight 9 are (11223), together with all forms

(111222), (111123), (1111122), (1111113), (11111112), (1 $\cdots$ 1)

derived from it by partition. Hence if (11223) represents all integers $<k$, the same is true of the remaining six forms. This was verified in the article just cited for $k=1400$. The fact that (11223) represents all integers $<40{,}000$ was proved in 1929,

but is here first published. First, we* added each sum of two cubes to each double of a sum of two cubes to obtain a list of all integers $<40{,}000$ which are represented by (1122). From each number not in that table we subtracted the triples of small cubes until we found a difference which is not in the table.

Another important result was proved: every integer >9580 and $\leqq 40{,}000$ is a sum of four cubes and the double of a cube. Since this is evidently true for all integers in our table of values of (1122), it sufficed to subtract doubles of small cubes from the numbers not in our table until we reached a difference which is a sum of four or fewer cubes. As a corollary, all integers between 9580 and 40,000 are sums of six cubes. This fact is an essential part of the proof that all integers are sums of nine cubes. That fact had been deduced in 1903 by von Sterneck [39], whose table was made by the successive additions of cubes, until sums of six were found. Any error by that method would introduce extremely many later errors. Our method avoided such a pyramiding of errors, and at any event gave an independent proof.

By means of our table for (1122), I verified also that each of the forms (11224), (11122), and (11222) represents all integers $<40{,}000$. Hence all the theorems in the Monthly hold to 40,000.

We shall not quote [40] the interesting empirical theorems for $n=4$.

For $n=5$, the order exceeds 7. The only two forms of order 8 which represent every integer $\leqq 7007$ are

$$(1123468,12), \quad (1123457,14).$$

Next, (1123469, 11) and (1123458, 13) represent all integers $\leqq 7007$ except 1931 and 3137, respectively. By partitioning a single coefficient of these four forms, we obtain 58 distinct forms of order 9 which represent all integers $\leqq 7007$. Consider $(11 \cdots, 10, 10)$ where the suppressed coefficients are 12345, 22245, 22335, or 22344; the first two represent all integers <3127, and the last two all $\leqq 2000$.

For $n=6$, R. C. Shook proved that the order exceeds 9, and obtained 22 forms of order 10 which represent all integers

* Half of the work was done at Chicago, mostly by assistants, and half by J. S. Turner of Ames, Iowa.

$\leqq 4^6 = 4{,}096$. They include (112457, 10, x, y, 17) for x, $y = 10$, 16; 11, 15; 12, 14; 13, 13.

For $n = 7$, K. C. Yang proved that the order exceeds 10, and verified that (11234579, 10, 15, 26, 60) of order 12 represents all integers $\leqq 16384$. M. E. Mauch obtained various general results for $n = 7$ and verified that (1123469, 13, 17, 41, 46) of order 11 represents all integers $\leqq 22{,}000$.

10. *Sums of Values of a Polynomial.* The early Greeks defined polygonal numbers. For example, the triangular numbers 1, $1+2$, $1+2+3, \cdots$ count the number of touching equal spheres arranged in the form of a triangle. The xth polygonal number of order $m+2$ (or $m+2$ sides) is

$$p_{m+2}(x) = x + m(x^2 - x)/2, \tag{1}$$

whence the xth triangular number is $p_3(x) = x(x+1)/2$. In 1636, Fermat [41] stated that he was the first to discover the beautiful theorem that every integer $\geqq 0$ is a sum of 3 triangular numbers $\geqq 0, \cdots$, and is a sum of m polygonal numbers $\geqq 0$ of order m.

Waring [42] conjectured that every positive integer P "of the proper form" is a sum of a finite number of values of a polynomial $f(x)$, where P is a multiple of 3 if $f(x) = 3x^4 + 6x^3 + 24$.

Gauss [43] proved that every integer $n = 8m+3$ is a sum of three odd squares $(2a_i+1)^2$, whence $m = \sum(a_i^2 + a_i)/2$, so that every integer $m \geqq 0$ is a sum of three triangular numbers. The number of ways m can be so decomposed depends in a definite manner on the prime factors of n and the number of classes of binary quadratic forms of determinant $-n$.

Barlow [44] noted that N is a sum of $m+2$ polygonal numbers (1) for $x = x_1, \cdots, x_{m+2}$ if and only if

$$8mN + (m+2)(m-2)^2 = \sum_{i=1}^{m+2} (2mx_i - m + 2)^2.$$

Cauchy [45] was the first to publish a proof of Fermat's assertion that every integer $A \geqq 0$ is a sum of $m+2$ polygonal p_{m+2}, all but four of which may be taken to be 0 or 1. In the simplified proof by Legendre [46], the case $m = 3$ is not presupposed, as was done by Cauchy. Moreover, Legendre proved that every integer $> 28(m-2)^3$ is a sum of four p_m if m is odd;

while, for m even, every integer $>7(m-2)^3$ is a sum of five p_m, one of which is 0 or 1.

Liouville [47] denoted triangular numbers by Δ's and proved that

$$\Delta+\Delta'+c\Delta'', \ (c=1,2,4,5), \quad \Delta+2\Delta'+d\Delta'', \ (d=2,3,4),$$

represent all integers $\geqq 0$, while no further linear combinations of three Δ's with positive integral coefficients represent all integers.

S. Réalis [48] proved that every integer is a sum of four of the numbers $(3z^2\pm z)/2$, and also a sum of four of the form $2z^2\pm z$, that is, p_5 and p_6 extended to negative arguments. [But three suffice, §11.]

Maillet [49] proved that if α and β are relatively prime odd integers, $\alpha>0$, every integer A exceeding a certain limit (depending on α, β) is a sum* of four integers $\geqq 0$ of the form $(\alpha x^2+\beta x)/2$. Next let α and β be both even and let $\alpha/2$ and $\beta/2$ be relatively prime. Then every sufficiently large odd integer is a sum of four integers $(\alpha x^2+\beta x)/2$ if $\alpha/2+\beta/2$ is odd.

Let α and β be relatively prime odd integers. According as the g. c. d. of $(\alpha+\beta)/2$ and 6 is 1, 2, 3 or 6, take $\rho=0$, 1, 2, 3, 4, or 5; $\rho=0$, 2, or 4; $\rho=0$ or 3; $\rho=0$, respectively. Then every sufficiently large integer $\equiv\rho \pmod 6$ is a sum of 53 values of $\alpha x^4/2+\beta x^2/2$ for integers $x\geqq 0$. He gave similar theorems when α and β are even.

Maillet [50] proved that, if a polynomial $f(x)$ is an integer $\geqq 0$ for every integer $x\geqq\mu$, then every integer exceeding a fixed limit, depending on the coefficients of $f(x)$, is a sum of at most n positive values of $f(x)$ and a limited number of 1's, where $n=6$, 12, 96, or 192, according as the degree of f is 2, 3, 4, or 5, respectively. Every integer $\geqq 19272$ is a sum of at most 12 pyramidal numbers $(x^3-x)/6$.

Kamke [51] proved the following precise formulation of Waring's conjecture. Let $f(x)$ be a polynomial of degree $n\geqq 2$ having rational coefficients, that of x^n being >0, such that $f(x)$ is an integer $\geqq 0$ for every integer $x\geqq 0$; then every integer $\geqq 0$ is a sum of a limited number of 1's and a limited number of values of $f(x)$ for integers $x\geqq 0$. The proof rests on the following

* We can assign an inferior limit to A, such that this decomposition can be made in any assigned number of ways.

interesting theorem on the simultaneous decomposition of n positive integers into powers of the same positive integers. If n is a given integer $\geqq 2$, there exist an integer $N=N(n)>0$, an integer $A>0$, and positive numbers i_1, i_ν, J_ν satisfying $0<i_\nu<J_\nu$, $(\nu=2, \cdots, n)$, such that for all integers $l_1, \cdots, l_n$ divisible by A, satisfying the inequalities

$$l_1 > i_1, \qquad i_\nu l_1^\nu < l_\nu < J_\nu l_1^\nu, \qquad (\nu = 2, \cdots, n),$$

the n equations

$$l_\nu = x_1^\nu + \cdots + x_N^\nu, \qquad (\nu = 1, \cdots, n),$$

are solvable simultaneously in integers $x_1, \cdots, x_N$.

Kamke [52] proved the following generalization of his preceding result. If $f(x)=a_1x^k+\cdots+a_k$ is an integer for every integer x, if $a_k>0$, and if T is the g. c. d. of the values of $f(x)$ for all integers x, then for all sufficiently large integers Z, TZ is a sum of a limited number of values $\geqq 0$ of $f(x)$. When $a_k=0$, it holds if T is the g. c. d. of all values of $f(x)-a_k$, and the TZ are replaced by the numbers of a suitably chosen class of residues modulo T.

Kamke's [51] existence theorem has been proved by the method of Hardy and Littlewood by Winogradow [53] and Landau [54].

The writer [55] proved that the only quadratic functions of x which are integers $\geqq 0$ for every integer $x\geqq 0$ and which take the values 0 and 1 for certain integers $x\geqq 0$ (necessary for a universal theorem) are the functions derived from the polygonal number (1) by replacing x by $x-k$ or $k-x$, where k is an integer $\geqq 0$. He found the least l such that every integer $\geqq 0$ is a sum of l values of such a function.

Consider any quadratic function $q(x)$ whose value is an integer $\geqq 0$ for every integer $x\geqq 0$. The problem to find the least e_s such that every positive integer is a sum of s values of $q(x)$ for integers $x\geqq 0$ and e_s terms 0 or 1 has been treated at length [56].

L. W. Griffiths [57] found all functions $a_1P_1+\cdots+a_nP_n$ which are universal, that is, represent all integers $\geqq 0$, where the P's are polygonal numbers of order $m+2$ and the a's are positive integers whose sum is $\leqq m+2$. She solved [58] the like

problem when P is $p_{m+2}(1-x)$, and partially the problem when P is the function (1) for $x=0, \pm 1, \pm 2, \cdots$.

Every integer is a sum of a square and two triangular numbers, and vice versa [59].

Dickson [60] proved that every sufficiently large integer is a sum of nine values of $x+k(x^3-x)/6$ for integers $x \geqq 0$, when k is prime to 3. This is true of all integers if k is 1 or 2.

James [61] proved that all sufficiently large integers are sums of eight pyramidal numbers $(x^3-x)/6$.

11. *Sums of Values of Two or More Polynomials.* For this new problem we have only examples.

Consider (1) and the extended polygonal number $e_{m+2}(x)$ of order $m+2$ which is derived from (1) by replacing x by $-x$. Both p and e are integers $\geqq 0$ for every integer $x \geqq 0$.

By a standard theorem, every positive integer of the form $8n+3$ is a sum of three squares, each of which is evidently odd. Any odd is of the form $4x \pm 1$. Hence $8n+3=\sum(4x \pm 1)^2$ for three integers $x \geqq 0$. Hence $n=\sum(2x^2 \pm x)$. For the upper sign, the summand is $e_6(x)$; for the lower sign, it is $p_6(x)$. Hence *every integer* $\geqq 0$ *is a sum of three numbers each of which is a value of one or the other of* $p_6(x)$, $e_6(x)$.

Similar theorems hold [62] for every m.

From $8n+3=\sum(2y+1)^2$, we see that every positive integer n is a sum of three triangular numbers $p_2(y)=(y^2+y)/2$. According as $y=2x$ or $y=2x-1$, $p_2(y)=e_6(x)$ or $p_6(x)$. This gives another proof of the theorem in italics. This method is applicable to any polynomial with fractional coefficients. For example, from $p_5(y)$ we get the pair $6x^2-x$, $6x^2+5x+1$. Hence every integer is a sum of five values of these two functions for integers $x \geqq 0$. Taking x to be $3y$, $3y+1$, $3y+2$ in turn in the last theorem of §10, with $k=1$, we see that every integer is a sum of nine values of

$$9y^3 + 2y,\ 9y^3 + 9y^2 + 5y + 1,\ 9y^3 + 18y^2 + 14y + 4$$

for integers $y \geqq 0$.

12. *Universal Sums of a Quadratic Form.* A binary quadratic form which represents unity is evidently equivalent to $B(x, y) = x^2+gxy+hy^2$, $g=0$ or 1. The case of sums of four or more values of B is trivial since every integer is a sum of four squares.

The writer [63] proved that every positive integer is a sum of two values of $B(x, y)$ for integers $x \geqq 0$, $y \geqq 0$ if and only if $h = 1$, 2, or 3; and is a sum of three values of B if and only if $h = 1, \cdots, 7$.

Asymptotic expressions have been found [64] for the number of representations of n by $N(b_1) + \cdots + N(b_r)$, where $N(b_i)$ is the norm of an ideal b_i of a class K_i of ideals of the quadratic field defined by $(D_i)^{1/2}$.

Bibliography

1. *Meditationes Algebraicae*, Cambridge, 1770, pp. 204–205; 3d ed., 1782, pp. 349–350.

2. Lectures at the Collège de France; printed in V. A. Lebesgue's *Exercices d'Analyse Numérique*, Paris, 1859, pp. 112–115.

3. Association Française Avancement des Sciences, vol. 24, II (1895), pp. 242–247.

4. Göttinger Nachrichten, 1909, pp. 17–36; Mathematische Annalen, vol. 67 (1909), pp. 281–300. See Schmidt, ibid., vol. 74 (1913), pp. 271–274.

5. Mathematische Annalen, vol. 65 (1908), pp. 424–427.

6. Mathematische Annalen, vol. 64 (1907), pp. 561–566.

7. Mathematische Annalen, vol. 66 (1909), p. 105.

8. Mathematische Annalen, vol. 66 (1909), pp. 95–101. A case overlooked was treated unsuccessfully by Bachmann and Lejneek, but the gap was finally filled by Kempner [18]. The proof was much simplified by Dickson, Transactions of this Society, vol. 30 (1928), pp. 1–18.

9. Mathematische Annalen, vol. 66 (1909), pp. 102–105; Landau, *Handbuch der Lehre von der Verteilung der Primzahlen*, vol. I, 1909, pp. 555–559. For numbers prime to 10 (without the theory of primes), Archiv der Mathematik und Physik, (3), vol. 18 (1911), pp. 248–252.

10. Mathematische Annalen, vol. 67 (1909), pp. 61–75.

11. Mathematische Annalen, vol. 67 (1909), pp. 301–305. See Hurwitz, ibid., vol. 73 (1912), pp. 173–176.

12. Arkiv för Matematik, Astronomi, Fysik, vol. 6 (1910–11), No. 32, No. 39. Summary in French in Mathematische Annalen, vol. 72 (1912), pp. 145–152.

13. Mathematische Zeitschrift, vol. 31 (1929), pp. 149–150.

14. Mathematische Annalen, vol. 72 (1912), pp. 153–156.

15. Sitzungsberichte Akademie Wissenschaften, Berlin, 1912, pp. 666–670.

16. Arkiv för Matematik, Astronomi, Fysik, vol. 11 (1916), No. 25; vol. 13 (1919), No. 25.

17. Bulletin Société Mathématique de France, vol. 36 (1908), pp. 69–77.

18. *Über das Waringsche Problem und einige Verallgemeinerungen*, Dissertation, Göttingen, 1912. Extract in Mathematische Annalen, vol. 72 (1912), p. 387.

19. *Beiträge zum Waringschen Problem*, Dissertation, Göttingen, 1913.

20. Extensions by Post, Dissertation, Halle, 1920, and Arndt, Dissertation, Göttingen, 1925.

21. Proceedings London Mathematical Society, (2), vol. 16 (1917), p. 130.

22. Quarterly Journal of Mathematics, vol. 48 (1919), pp. 272–293. See Hardy, *Some Famous Problems of the Theory of Numbers, and in particular Waring's Problem*, Oxford, 1920.

23. I. Göttinger Nachrichten, 1920, pp. 33–54.
II. Mathematische Zeitschrift, vol. 9 (1921), pp. 14–27.
III. Acta Mathematica, 1922, pp. 1–70 (sum of primes).
IV. Mathematische Zeitschrift, vol. 12 (1922), pp. 161–188.
V. Proceedings London Mathematical Society, (2), vol. 22 (1923), pp. 46–56 (Goldbach's problem).
VI. Mathematische Zeitschrift, vol. 23 (1925), pp. 1–37.
VII. Proceedings London Mathematical Society, (2), vol. 28 (1927), pp. 518–542.

24. Göttinger Nachrichten, 1921, pp. 88–92; Mathematische Zeitschrift, vol. 12 (1922), pp. 219–247; Proceedings London Mathematical Society, (2), vol. 25 (1926), pp. 484–485. *Vorlesungen über Zahlentheorie*, I, 1927, pp. 235–360 (pp. 276–303 may be replaced by his simpler proof in Mathematische Zeitschrift, vol. 32 (1930), pp. 699–702). See papers [53], [54].

25. Göttinger Nachrichten, 1921, pp. 189–192; 1922.

26. Mathematische Annalen, vol. 105 (1931), pp. 637–652.

27. Mathematische Zeitschrift, vol. 9 (1921), pp. 28–34.

28. British Association Report, vol. 90 (1922), pp. 16–24.

29. Journal London Mathematical Society, vol. 1 (1926), pp. 244–245.

30. Ibid., pp. 72–74. Improved results by Pillai, ibid., vol. 3 (1928), pp. 56–61, 83.

31. Ibid., vol. 5 (1930), pp. 155–158.

32. Ibid., vol. 6 (1931), pp. 194–197.

33. Proceedings London Mathematical Society, 1933–34.

34. American Journal Mathematics, vol. 49 (1927), pp. 241–250.

35. Transactions of this Society, vol. 30 (1928), pp. 1–18.

36. Dissertation, Chicago, 1928.

37. Proceedings London Mathematical Society, 1933–34.

38. Dickson, American Mathematical Monthly, vol. 34 (1927), pp. 177–183.

39. Sitzungsberichte, Akademie der Wissenschaften, Wien, vol. 112, II a (1903), pp. 1627–1666.

40. Dickson, this Bulletin, vol. 33 (1927), pp. 319–325.

41. Oeuvres, vol. I, p. 305; vol. II, p. 65. French translation, III, p. 252, p. 287.

42. *Meditationes Algebraicae*, 1770, pp. 204–205; 3d ed., 1782, pp. 349–350.

43. *Disquisitiones Arithmeticae*, 1801, §293; Werke, I, 1863, p. 348.

44. *Theory of Numbers*, 1811, p. 219.

45. Mémoires Scientifiques Mathématiques et Physiques de l'Institut de France, (1), vol. 14 (1813–15), pp. 177–220; same in *Exercices de Mathématiques*, vol. I, 1826, pp. 265–296; Oeuvres de Cauchy, (2), vol. VI, pp. 320–353; Journal des Mines, vol. 38 (1815), p. 395. The proof by Pepin, Atti Accademia Pontifica Nuovi Lincei, vol. 48 (1892–93), pp. 119–131, holds

when $A > 110m$, and was verified by a table for all $A < 120m+16$. The simpler proof by Dickson [56] requires a table only to $48m+21$.

46. *Théorie des Nombres*, first supplement, 1816, to the second edition, 1808, pp. 13–27; 3d ed., 1830, vol. I, p. 218; vol. II, p. 340; German translation by Maser, vol. 2, p. 332.

47. Journal de Mathématiques, (2), vol. 7 (1862), p. 407; vol. 8 (1863), p. 73.

48. Nouvelles Annales Mathématiques, (2), vol. 12 (1873), p. 212; Nouvelle Correspondance Mathématique, vol. 4 (1878), pp. 27–30.

49. Bulletin Société Mathématique de France, vol. 23 (1895), pp. 40–49.

50. Journal de Mathématiques, (5), vol. 2 (1896), pp. 363–380.

51. Mathematische Annalen, vol. 83 (1921), pp. 85–112. In Journal für Mathematik, vol. 152 (1923), pp. 30–32, he evaluated A when we employ only the values l and n of ν.

52. Mathematische Zeitschrift, vol. 15 (1922), pp. 188–194; vol. 19 (1924), p. 263.

53. Recueil Société Mathématique Moscow, vol. 31 (1924), pp. 490–507. Akademia Nauk, Bulletin, 1928, pp. 393–400; pp. 401–414. (In Russian.)

54. Acta Mathematica, vol. 48 (1926), pp. 217–253; Mathematische Zeitschrift, vol. 12 (1922), pp. 219–247; vol. 31 (1929), pp. 149–150, 319–338.

55. This Bulletin, vol. 33 (1927), pp. 713–720; vol. 34 (1928), pp. 63–72, 205–217.

56. Dickson, Journal de Mathématiques, (9), vol. 7 (1898), pp. 328–336; American Journal Mathematics, vol. 50 (1928), pp. 1–48; Pall, ibid., vol. 54 (1932), pp. 66–78; Transactions of this Society, vol. 34 (1932), pp. 98–125; this Bulletin, vol. 37 (1931), pp. 267–270; Quarterly Journal Mathematics, vol. 2 (1931), pp. 136–143.

57. Annals of Mathematics, (2), vol. 31 (1930), pp. 1–12.

58. Abstracts in this Bulletin, vol. 35 (1929), p. 767; vol. 36 (1930), p. 806.

59. Nouvelles Annales Mathématiques, (2), vol. 11 (1872), pp. 95–6, 516–519; (2), vol. 12 (1873), p. 217.

60. Transactions of this Society, 1933–34.

61. Mathematische Annalen, 1933.

62. Dickson, Journal de Mathématiques, (9), vol. 7 (1928), p. 327.

63. Journal de Mathématiques, (9), vol. 7 (1898), pp. 319–326.

64. Peterson and Rademacher, Abhandlungen, Mathematisches Seminar Hamburg Universität, vol. 4 (1925), pp. 267–296; vol. 5, pp. 40–44.

THE UNIVERSITY OF CHICAGO

ABSTRACT IDEAL THEORY*

BY OYSTEIN ORE

1. *Introduction.* Abstract ideal theory is a branch of algebra which has come into prominence only in recent years; its importance for algebraic problems and also for branches of mathematics outside of algebra proper has however been increasingly recognized; it seems established that the ideals, corresponding in many ways to the normal subgroups in the theory of groups, are the most convenient building stones in a large number of algebraic structures.

In the following I have tried to give a survey of the most important problems and results, but it should be realized that an account of this kind must necessarily be incomplete, since the field is too wide and too diverse to be covered in a single lecture. To limit the subject, only commutative ideal theory will be considered; demonstrations have been omitted, although reluctantly, since an occasional proof will often clarify more than any explanation the tools and working methods of a theory. Another difficulty in this case lies in the fact that the theory itself to a large extent is still in an evolutionary stage and has not reached the harmonious form it will probably assume later on. Only for domains in which the finite chain condition holds does it seem to have arrived at some degree of perfection. Historically several of the fundamental ideas can be traced to the work of Dedekind, Kronecker, and Lasker, but the main contributions to the theory have been made in the last ten years by E. Noether, Krull, van der Waerden, Henzelt, Grell, Stiemke, and others.

The first part of the paper contains some of the main properties of ideals, operations on ideals, quotient rings, and isomorphisms; then follows a discussion of prime and primary ideals and the consequences of the finite divisor chain condition. The four principal ideal representations by E. Noether are treated fairly completely, and to illustrate these abstract de-

* Symposium lecture delivered at the meeting of the Society in New York, April 14, 1933.

composition theorems I have given some of their applications to algebraic manifolds and the theorems of Hilbert and M. Noether. The rings in which every ideal is the product of prime ideal powers, that is, the integrally closed rings, are mentioned next; in the last part I have given some results for rings in which the chain condition does not hold and also discussed some of the novel ideas of Krull on a more topological introduction and definition of ideals by means of absolute values.

2. *Ideals and Operations on Ideals.* In the following we shall study the structural properties of an abstract ring R; we assume that R has the ordinary ring properties, that is, it is an abelian group with respect to addition, and that multiplication is associative, distributive, and commutative. An *ideal* $\mathfrak{a}$ in R is a subring of R with the following properties:

(a) When α and β are elements of R belonging to $\mathfrak{a}$, then $\alpha \pm \beta$ belongs to $\mathfrak{a}$.

(b) For an arbitrary element λ in R and α in $\mathfrak{a}$ the product $\lambda\alpha$ belongs to $\mathfrak{a}$.

This definition is due to Dedekind. An important class of ideals are the *principal ideals* (α) generated by a single element α in R; (α) must contain all sums $\alpha+\alpha \cdots$, and all multipla $\lambda\alpha$ such that every element can be written in the form $\alpha' = \lambda\alpha + n\alpha$, where n is a rational integer; when R contains a unit element ϵ we obtain the simpler representation $\alpha' = \lambda\alpha$. Another class of ideals are those having an *ideal basis*, $\mathfrak{a} = (\alpha_1, \cdots, \alpha_r)$; every element in $\mathfrak{a}$ can then be represented in the form

$$\alpha' = \lambda_2\alpha_1 + \cdots + \lambda_r\alpha_r + n_1\alpha_1 + \cdots + n_r\alpha_r,$$

where the λ_i belong to R and the n_i are rational integers. We suppose in the following that R does not consist only of the zero element; every ring then has at least two ideals, the *zero ideal* $\mathfrak{n} = (0)$ and the *unit ideal* R; when R contains a unit element then $R = (\epsilon)$ is a principal ideal.

The fact that an element α belongs to an ideal $\mathfrak{a}$ is usually expressed by a congruence $\alpha \equiv 0 \pmod{\mathfrak{a}}$; the more general congruence $\alpha \equiv \beta \pmod{\mathfrak{a}}$ denotes that $\alpha - \beta$ belongs to $\mathfrak{a}$. When all elements of an ideal $\mathfrak{b}$ belong to $\mathfrak{a}$, we write $\mathfrak{b} \equiv 0 \pmod{\mathfrak{a}}$ or $\mathfrak{a} > \mathfrak{b}$ and say that $\mathfrak{b}$ is *contained in* $\mathfrak{a}$ or $\mathfrak{a}$ is a *divisor of* $\mathfrak{b}$. It is

obvious that the zero ideal $\mathfrak{n}$ is divisible by all ideals and that R divides all ideals.

The application of ideals to algebraic problems is based upon the existence of various processes by means of which one can derive new ideals from a set of given ideals. There are four fundamental ways of combining two given ideals $\mathfrak{a}$ and $\mathfrak{b}$ into new ideals, and since they will be used constantly in the following, their definitions and simplest properties shall be mentioned here.

(i). *The cross-cut* or *greatest common divisor* $\mathfrak{d}=(\mathfrak{a}, \mathfrak{b})$ of two ideals consists of all elements of the form $\delta=\alpha+\beta$, where α belongs to $\mathfrak{a}$ and β to $\mathfrak{b}$; it has the properties

$$\mathfrak{a} \equiv 0, \ \mathfrak{b} \equiv 0 \pmod{\mathfrak{d}}, \tag{1}$$

and $\mathfrak{d} \equiv 0 \pmod{\mathfrak{d}_1}$ for all other ideals $\mathfrak{d}_1$ for which (1) holds.

(ii). The *union* or *least common multiplum* $\mathfrak{m}=[\mathfrak{a}, \mathfrak{b}]$ consists of all elements contained both in $\mathfrak{a}$ and $\mathfrak{b}$; one has

$$\mathfrak{m} \equiv 0 \pmod{\mathfrak{a}}, \qquad \mathfrak{m} \equiv 0 \pmod{\mathfrak{b}}, \tag{2}$$

and $\mathfrak{m} \equiv 0 \pmod{\mathfrak{m}}$ for all other ideals $\mathfrak{m}$ satisfying (2).

(iii). The *product* $\mathfrak{ab}$ consists of all elements $\gamma=\Sigma_1' \alpha\beta$ representable as the sum of products $\alpha\beta$, where α belongs to $\mathfrak{a}$ and β to $\mathfrak{b}$; the multiplication of ideals is obviously associative and commutative. When $\mathfrak{c}=\mathfrak{ab}$, we say that $\mathfrak{a}$ is a *factor* of $\mathfrak{c}$; it is seen that then $\mathfrak{c} \equiv 0 \pmod{\mathfrak{a}}$ so that $\mathfrak{a}$ is also a divisor of $\mathfrak{c}$, but not conversely. From the former definitions one concludes

$$\mathfrak{a}(\mathfrak{b}, \mathfrak{c}) = (\mathfrak{ab}, \mathfrak{ac}), \qquad \mathfrak{ab} \equiv 0 \pmod{[\mathfrak{a}, \mathfrak{b}]},$$

and it can also be shown that

$$[\mathfrak{a}, \mathfrak{b}](\mathfrak{a}, \mathfrak{b}) \equiv 0 \pmod{\mathfrak{ab}}.$$

(iv). The *ideal quotient* $\mathfrak{a}:\mathfrak{b}$ is the set of all elements γ in R such that $\gamma\beta$ belongs to $\mathfrak{a}$ for any element β in $\mathfrak{b}$. If $\mathfrak{c}=\mathfrak{a}:\mathfrak{b}$, then

$$\mathfrak{cb} \equiv 0 \pmod{\mathfrak{a}}, \tag{3}$$

and $\mathfrak{c}_1 \equiv 0 \pmod{\mathfrak{c}}$ for any other ideal $\mathfrak{c}_1$ satisfying (3). Among the most important properties of the ideal quotient are the following:

$$\mathfrak{a}:\mathfrak{b} = \mathfrak{a}:(\mathfrak{a}, \mathfrak{b}),$$

$$[\mathfrak{a}, \mathfrak{b}]:\mathfrak{c} = [\mathfrak{a}:\mathfrak{c}, \mathfrak{b}:\mathfrak{c}],$$

$$(\mathfrak{a}:\mathfrak{b}):\mathfrak{c} = (\mathfrak{a}:\mathfrak{c}):\mathfrak{b} = \mathfrak{a}:\mathfrak{bc}.$$

All the operations introduced here can of course be extended to the case of an arbitrary finite number of ideals.

3. *Quotient Rings.* Let $\mathfrak{a}$ be an ideal and let us divide all of the elements of R into classes (mod $\mathfrak{a}$) by stipulating that two elements α and α_1 belong to the same class $\bar{\alpha}$ if and only if $\alpha \equiv \alpha_1 \pmod{\mathfrak{a}}$. The sum and product of two classes $\bar{\alpha}+\bar{\beta}$ and $\bar{\alpha}\bar{\beta}$ are then defined as the classes containing $\alpha+\beta$ and $\alpha\beta$; by this definition the set of all residue classes forms a commutative ring, which we denote by $R/\mathfrak{a}$ and call the *quotient ring* or *ring of residue classes* of R with respect to $\mathfrak{a}$.

At this point it becomes clearer why we have chosen, from the beginning, to study ideals among the various possible types of subrings in R: *the ideals are the only subrings for which the quotient ring exists.* The ideals correspond to the normal subgroups in the group theory, these being the only sub-groups for which a quotient group exists. This analogy between ideals and normal sub-groups pervades the whole theory of ideals and we shall see several instances of it in the following; let us mention as an example: Every ideal $\bar{\mathfrak{c}}$ in the quotient ring $R/\mathfrak{a}$ is an ideal $\mathfrak{c}$ in R containing $\mathfrak{a}$ when considered as a set of elements in R and conversely.

This analogy appears most clearly in the study of homeomorphisms of rings. We say that R corresponds *homeomorphically* to a ring R' when to every element α in R there corresponds a unique α' in R' such that $\alpha \pm \beta$ corresponds to $\alpha' \pm \beta'$ and $\alpha\beta$ to $\alpha'\beta'$, while conversely there exists at least one α to each α'. When the homeomorphism is a one-to-one correspondence we say that R and R' are *isomorphic*; a homeomorphism is usually denoted by $R \sim R'$ and an isomorphism by $R \cong R'$.

In the case of a homeomorphism between R and R', one finds that all elements in R corresponding to the zero element $0'$ in R' form an ideal $\mathfrak{a}$ in R since from $\alpha \to 0'$, $\alpha_1 \to 0'$ follows $\alpha \pm \alpha' \to 0'$ and $\lambda\alpha \to 0'$ for an arbitrary λ in R. All the elements in an arbitrary residue class $\bar{\gamma}$ (mod $\mathfrak{a}$) in R must then correspond to the same element γ' in R and this correspondence is easily seen to be an isomorphism. This gives the fundamental theorem on the homeomorphisms of rings. *Every ring homeomorphic to R is isomorphic to a residue ring $R/\mathfrak{a}$, where $\mathfrak{a}$ is an ideal in R, and every such residue ring is homeomorphic to R.* The second part of the theorem is evident.

It should be observed that the construction of residue rings and homeomorphic rings is a very familiar process in algebra and even in analysis. Let us consider, as an example, the construction of the field of real numbers K from the field k of rational numbers. One forms the ring R of all convergent sequences $\alpha=(a_0, a_1, \cdots)$, where the a_i are rational numbers. The set of all sequences having the limit zero is an ideal σ in R. Next one proceeds to identify all zero sequences, that is, one constructs the quotient ring R/σ and after showing that this ring is a field, K is defined by the relation $K \cong R/\sigma$. The same process can obviously be applied in any ring in which an absolute value or the notion of a limit can be defined.

I also mention that from the definitions (i) and (ii) one can derive a second fundamental theorem on the isomorphisms of rings expressible in the formula

$$(\mathfrak{a}, \mathfrak{b})/\mathfrak{a} \cong \mathfrak{b}/[\mathfrak{a}, \mathfrak{b}]. \tag{4}$$

4. *Prime Ideals.* A natural generalization of prime ideals in the theory of algebraic numbers would be to define a *prime ideal* $\mathfrak{p}$ as *an ideal without divisors*, that is, there shall exist no ideal except $\mathfrak{p}$ itself and R containing it. A prime ideal might then also be called a *maximal ideal* and according to a remark in §3 the quotient ring $R/\mathfrak{p}$ must be simple, that is, have no ideals different from the ring itself and the zero ideal. A simple analysis yields the following result.* *A simple ring is a field or a ring isomorphic to the finite ring*

$$R_p = 0, p, 2p, \cdots, (p-1)p \pmod{p^2}. \tag{5}$$

This definition of a prime ideal was actually used by Sono, one of the earliest writers on abstract ideal theory. However, the applications of the theory have shown that this definition is too limited and that a more satisfactory definition is obtained by using another property of prime numbers. A prime ideal $\mathfrak{p}$ is an ideal such that from $\alpha\beta \equiv 0 \pmod{\mathfrak{p}}$ follows $\alpha \equiv 0$ or $\beta \equiv 0 \pmod{\mathfrak{p}}$. One may also state this definition in an equivalent form: *from an ideal relation* $\mathfrak{ba} \equiv 0 \pmod{\mathfrak{p}}$ *follows* $\mathfrak{b} \equiv 0$ or $\mathfrak{a} \equiv 0 \pmod{\mathfrak{p}}$, or finally as a property of the residue ring: *a prime ideal is an ideal such that the residue ring* $R/\mathfrak{p}$ *is a domain of integrity, that is, has no divisors of zero.*

* Sono, Memoirs of the College of Science, Kyoto, vol. 2 (1917).

In the following we shall adopt the last definition of a prime; it is easily seen that it is not equivalent to the former since the zero ideal $\mathfrak{n}$ in (5) is a prime ideal by the first definition, $R_p/\mathfrak{n} \cong R_p$ being simple and yet $R_p^2 = \mathfrak{n}$.

A very characteristic, but at the first glance somewhat surprising, property of our prime ideals is that they may have divisors. As an example let us consider the ring of all polynomials in two variables x and y with coefficients in a field K. The ideal $\mathfrak{p}_1 = (y)$ is seen to be prime and to consist of all polynomials divisible by y; the residue ring is the set of all polynomials in x. The ideal $\mathfrak{p}_2 = (x, y)$ is a prime ideal dividing $\mathfrak{p}_1$; it consists of all polynomials without constant term, and the residue ring is K. We shall see later that by certain geometrical applications, if there exists a chain

$$\mathfrak{p}_0 < \mathfrak{p}_1 < \cdots < \mathfrak{p}_r < R$$

of prime ideal divisors of a prime ideal $\mathfrak{p}_0$ such that no further prime ideal can be inserted in the chain, then r coincides with the dimension of the manifold corresponding to $\mathfrak{p}_0$.

5. *The Chain Condition.* We have introduced ideals in order to study the structure of a ring R and its subrings. In a special case, as for instance the ring of all integers in a finite algebraic field, we have the fundamental theorem that every ideal is uniquely representable as the product of prime ideals. For more general rings a result of this simple nature cannot be expected as examples show; as a matter of fact, the ideal theory for the most general rings is so complicated that it is at the present time not fully mastered. We shall therefore have to impose a condition on the rings we consider; this condition brings us closer to the case of ideals in algebraic fields, but it is at the same time sufficiently general to include the most important rings in algebra. We shall call it the *condition of finite divisor chains* or simply *the chain condition.*

Every chain of ideals $\mathfrak{a}_1 < \mathfrak{a}_2 < \cdots$, *each dividing the preceding, must break off after a finite number of terms.* One can also put the condition in the form, that if an infinite chain $\mathfrak{a}_1 \leqq \mathfrak{a}_2 \leqq \cdots$ exists, then (from a certain point on) all ideals must be equal. The chain condition corresponds to the condition usually imposed by the study of groups, that the series of composition shall be of finite length. The condition is obviously satisfied in

rings of integers in a finite algebraic field, since all ideals have only a finite number of divisors.

One may state the finite chain condition in a different manner, which is sometimes more useful for applications. *When the chain condition holds for the ideals in a ring R, then every ideal has a finite ideal basis and conversely.* Suppose that the chain condition holds and let $\mathfrak{a}$ be an ideal in R; when α_1 is an element of $\mathfrak{a}$, then $\mathfrak{a}=(\alpha_1)$ or there exists an element α_2 in $\mathfrak{a}$, but not in (α_1); then we have $\mathfrak{a}=(\alpha_1, \alpha_2)$ or there exists an α_3 in $\mathfrak{a}$ not contained in (α_1, α_2), and so on. This process must finally give us a basis for $\mathfrak{a}$ since

$$(\alpha_1) < (\alpha_1, \alpha_2) < (\alpha_1, \alpha_2, \alpha_3) < \cdots$$

breaks off. Let us suppose conversely that every ideal has a basis and let $\mathfrak{a}_1 \leqq \mathfrak{a}_2 \leqq \cdots$ be an infinite chain of ideals. The union of all ideals $\mathfrak{a}_i$ is an ideal $\mathfrak{b}=(\beta_1, \cdots, \beta_r)$ in R, and since each β_i belongs to an ideal $\mathfrak{a}_j$ and to all following ideals, one must have $\mathfrak{a}_k=\mathfrak{b}$ from a certain index k.

6. *Irreducible Ideals.* The idea of decomposition theorems for ideals in rings where unique decomposition into prime ideal factors does not exist, seems to have originated with Dedekind; Lasker* realized their importance for polynomial ideals and proved the existence of a decomposition into primary components; an account of Lasker's work can be found in Macaulay's tract: *Algebraic Theory of Modular Systems.* The first general investigation of the decomposition of ideals in rings (satisfying the chain condition) is due, however, to E. Noether;† in the following, we shall give the principal results which have been obtained.

An ideal $\mathfrak{a}$ is said to be *reducible* when it can be represented as the union $\mathfrak{a}=[\mathfrak{b}, \mathfrak{c}]$ of two proper divisors $\mathfrak{b}$ and $\mathfrak{c}$; $\mathfrak{a}$ is *irreducible* when no such representation exists. When $\mathfrak{a}$ is reducible, the divisors $\mathfrak{b}$ and $\mathfrak{c}$ may be decomposed further, and from the chain condition, it follows that one can finally obtain a representation

$$\mathfrak{a} = [\mathfrak{a}_1, \mathfrak{a}_2, \cdots, \mathfrak{a}_r], \tag{6}$$

* Mathematische Annalen, vol. 60 (1905).

† Mathematische Annalen, vol. 83 (1921).

where all $\mathfrak{a}_i$ are irreducible. It may occur that some of the ideals $\mathfrak{a}_i$ in the representation (6) are superfluous, $\mathfrak{a}_i$ being a divisor of the union of the remaining ideals; we shall suppose that all such ideals are omitted. It may also happen that a component can be replaced by a proper divisor. When all such reductions have been made, we say that (6) is a *reduced representation of* $\mathfrak{a}$ *by means of irreducible components.* This is the first decomposition given by E. Noether; it is usually not unique, but one can show: *The number of irreducible components in two different reduced representations is always the same.*

7. *Primary Ideals.* In the ring of rational integers the reduced representation (6) corresponds to the representation of an integer as the product or least common multiplum of prime powers. It may, therefore, be natural to seek a connection between the reduced representation and the prime ideals dividing $\mathfrak{a}$; this connection is obtained through the introduction of *primary ideals.* These ideals form a generalization and a substitute for the prime ideal powers in the general case and reduce to them in the case of unique prime ideal factorization. *An ideal* $\mathfrak{q}$ *is said to be primary if a congruence* $\alpha\beta \equiv 0 \pmod{\mathfrak{q}}$, *where* $\alpha \equiv 0 \pmod{\mathfrak{q}}$, *implies the existence of an exponent* k *such that* $\beta^k \equiv 0 \pmod{\mathfrak{q}}$. When $k=1$ for all elements in R, then $\mathfrak{q}$ is a prime ideal. To every primary ideal $\mathfrak{q}$ there exists, as for the prime powers, a unique prime ideal $\mathfrak{p}$ associated with it; this prime ideal is defined as the set of all elements π in R having the property that a power π^z belongs to $\mathfrak{q}$. A simple study shows that these elements actually form an ideal which must be prime. The definition of $\mathfrak{p}$ and the primary ideal $\mathfrak{q}$ shows that a congruence $\alpha\beta \equiv 0 \pmod{\mathfrak{q}}$, where $\alpha \equiv 0 \pmod{\mathfrak{q}}$, implies that β belongs to $\mathfrak{p}$.

Since we suppose that the chain condition holds, there exists a representation $\mathfrak{p} = (\beta_1, \cdots, \beta_2)$. A power of each β_i belongs to $\mathfrak{q}$ according to definition and a power of $\mathfrak{p}$ must consequently also be divisible by $\mathfrak{q}$. *To every primary ideal* $\mathfrak{q}$ *there belongs a unique prime such that*

$$\mathfrak{p}, \mathfrak{q} > \mathfrak{p}^r. \tag{7}$$

One might have used the relation (7) to define a primary ideal and its corresponding prime ideal. A primary ideal is sometimes defined by an ideal condition, namely, that the ideal congruence

$$\mathfrak{a}\mathfrak{b} \equiv 0 \pmod{\mathfrak{q}}, \quad \mathfrak{a} \not\equiv 0 \pmod{\mathfrak{q}},$$

shall imply $\mathfrak{b}^k \not\equiv 0 \pmod{\mathfrak{q}}$. An ideal having this property is said to be *strongly primary*; while the former definition yields *weakly primary* ideals. The two concepts are identical in rings in which the chain condition holds.

The study of the structure of primary ideals seems to be one of the most important problems in abstract ideal theory. Various authors have dealt with this problem directly or with questions which can be interpreted in this way. I shall only mention papers by Sono, Schmeidler, van der Waerden, and a series of contributions by Krull. The problem is equivalent to the investigation of *primary rings*, that is, the rings of residue classes of primary ideals. The definition of a primary shows that these rings have the characteristic property that a power of every divisor of zero must vanish.

8. *Decomposition Theorems.* We now return to the representation (6) of an ideal by means of irreducible components. It can be shown, that *every irreducible ideal is primary*; each component, consequently, has a characteristic prime ideal and it can be proved, furthermore, that not only is the number of irreducible components in any representation (6) of an ideal $\mathfrak{a}$ the same, but also the corresponding prime ideals must be the same and have the same multiplicity.

Another theorem on primary ideals is the following. *The union of primary ideals belonging to the same prime ideal is again primary, but the union of primary ideals belonging to different prime ideals is never primary.* We now unite all primary ideals in (6) belonging to the same prime ideal $\mathfrak{p}$ into a new primary ideal $\mathfrak{b}_\mathfrak{p}$, and call $\mathfrak{b}_\mathfrak{p}$ a *maximal primary component* of $\mathfrak{a}$. Our former results then yield this: *Every ideal* $\mathfrak{a}$ *can be represented as the union of maximal primary components*

$$\mathfrak{a} = [\mathfrak{b}_1, \mathfrak{b}_2, \cdots, \mathfrak{b}_s]; \tag{8}$$

the number of components and the corresponding prime ideals

$$\mathfrak{p}_1, \mathfrak{p}_2, \cdots, \mathfrak{p}_s \tag{9}$$

are the same for all such decompositions. The prime ideals (9) are all different and uniquely determined; we shall call them the *prime ideals belonging to* $\mathfrak{a}$.

E. Noether gives two other types of decomposition theorems;

the two last types are distinguished by the fact that the decompositions are unique.

An ideal $\mathfrak{b}$ is said to be *relatively prime* to $\mathfrak{a}$ if any relation $\mathfrak{b}c \equiv 0 \pmod{\mathfrak{a}}$ implies $c \equiv 0 \pmod{\mathfrak{b}}$. If we suppose $\mathfrak{a} < R$, then the necessary and sufficient condition that $\mathfrak{b}$ be relatively prime to $\mathfrak{a}$ is that no prime ideal belonging to $\mathfrak{b}$ is divisible by a prime ideal belonging to $\mathfrak{a}$. This criterion shows that the notion of being relatively prime is not mutual, that is, $\mathfrak{b}$ may be relatively prime to $\mathfrak{a}$, but not $\mathfrak{a}$ to $\mathfrak{b}$. Krull defines relatively prime by the equivalent condition $\mathfrak{a}:\mathfrak{b} = \mathfrak{a}$. An ideal $\mathfrak{a}$ is said to be *relatively-prime decomposable* when there exists no representation $\mathfrak{a} = [\mathfrak{b}, \mathfrak{c}]$, where $\mathfrak{b}$ and $\mathfrak{c}$ are mutually relatively prime. One can then show that *every ideal has a unique decomposition*,

$$\mathfrak{a} = [\mathfrak{c}_1, \mathfrak{c}_2, \cdots, \mathfrak{c}_t], \tag{10}$$

as the union of relatively-prime indecomposable components. The decomposition (10) can be obtained from the maximal primary decomposition (8) by joining all primary components into sets such that when a set contains a primary ideal with the corresponding prime ideal $\mathfrak{p}$, then it contains all primary ideals belonging to prime divisors or multipla of $\mathfrak{p}$. The uniqueness of the decomposition (10) also gives some information about the components in (8), and a certain class of maximal primary components, the so-called *isolated components*, can be shown to be uniquely determined.

The final decomposition theorem does not possess quite the same generality as the preceding since we shall have to suppose that *the ring contains a unit element* ϵ. On the other hand, this decomposition claims additional interest since it gives not only a representation as the union of ideal components, but also a product representation by means of the same ideals. Let $\mathfrak{a}$ and $\mathfrak{b}$ be ideals without common divisors, that is, $(\mathfrak{a}, \mathfrak{b}) = R = (\epsilon)$; this implies that $\mathfrak{a}$ and $\mathfrak{b}$ are mutually relatively prime, and the definitions in §2 show that in this case $\mathfrak{a}\mathfrak{b} = [\mathfrak{a}, \mathfrak{b}]$; *for ideals without common divisors the union is equal to the product.* An ideal is *direct indecomposable*, when no identity $\mathfrak{a} = [\mathfrak{b}, \mathfrak{c}]$ holds, where $\mathfrak{b}$ and $\mathfrak{c}$ are without common divisors. It can then be shown that *every ideal is uniquely representable as the union or product of direct indecomposable ideals,*

$$\mathfrak{a} = [\mathfrak{d}_1, \cdots, \mathfrak{d}_n] = \mathfrak{d}_1 \cdots \mathfrak{d}_n, \tag{11}$$

where each $\mathfrak{d}_i$ *has no common divisor with the union of the rest.*

The four decompositions (6), (8), (10), (11) give the main results on the structure of ideals; they have been given, as in the original paper by E. Noether, in order of decreasing decomposition, each being deducible from the preceding by joining some of its components into larger groups. Space does not permit a discussion of other properties of these decompositions, but it should be mentioned that Krull* has simplified the proofs of the main theorems through a systematic use of the ideal quotient. In another paper† Krull deduces the theory on an axiomatic basis, considering the ideals not as sets in rings, but as symbols having the operational properties defined in §2.‡

9. *Polynomial Ideals and Algebraic Manifolds.* To clarify the significance of this abstract theory it may be well to consider a concrete illustration; a suitable example of sufficient generality can be found in the theory of polynomial ideals and the related theory of algebraic manifolds. We shall suppose that K is an algebraically closed field and consider the ring $R(x_1, \cdots, x_n)$ of all polynomials in n variables with coefficients in K; according to a well known theorem by Hilbert the chain condition will hold in this ring since every ideal has a basis.

An *algebraic manifold* in the corresponding n-dimensional space S_n is the totality of points $\xi = (\xi_1, \cdots, \xi_n)$ with coordinates in K satisfying a set of algebraic relations

$$f_i(x) = f_i(x_1, \cdots, x_n) = 0, \quad (i = 1, 2, \cdots). \tag{12}$$

To every ideal $\mathfrak{a}$ in $R(x_1, \cdots, x_n)$ then corresponds a unique algebraic manifold $\mathfrak{M}(\mathfrak{a})$ defined as the totality of points at which all polynomials of $\mathfrak{a}$ vanish; it coincides with the manifold defined by the basis polynomials of $\mathfrak{a}$. Two different ideals may have the same manifold, for instance, $\mathfrak{M}(\mathfrak{a}) = \mathfrak{M}(\mathfrak{a}^n)$; when $\mathfrak{a} > \mathfrak{b}$, then obviously

$$\mathfrak{M}(\mathfrak{a}) \leqq \mathfrak{M}(\mathfrak{b}).$$

To every algebraic manifold $\mathfrak{M}$ corresponds conversely a unique

* Mathematische Annalen, vol. 90 (1923).

† Sitzungsberichte Erlangen, vol. 56 (1924).

‡ See also Sono, Memoirs of the College of Science, Kyoto, vol. 7 (1923).

ideal $\mathfrak{a}_{\mathfrak{M}}$ defined as the set of all polynomials vanishing in all points of $\mathfrak{M}$.

One easily finds that the cross-cut of two algebraic manifolds $\mathfrak{M}$ and $\mathfrak{N}$ is an algebraic manifold defined by $(\mathfrak{a}_{\mathfrak{M}}, \mathfrak{a}_{\mathfrak{N}})$, while the union is an algebraic manifold defined by the union $[\mathfrak{a}_{\mathfrak{M}}, \mathfrak{a}_{\mathfrak{N}}]$ or also by the product $\mathfrak{a}_{\mathfrak{M}}\mathfrak{a}_{\mathfrak{N}}$. We say, as usual, that an algebraic manifold is *irreducible* if it is not the union of two manifolds contained in it; one can then show that *the necessary and sufficient condition that a manifold be irreducible is that the corresponding ideal be a prime ideal; every prime ideal defines a unique irreducible algebraic manifold.*

This result illustrates clearly the occurrence of divisors of prime ideals as has already been mentioned in §4. For instance, in three-dimensional space any irreducible surface will correspond to a prime ideal $\mathfrak{p}_2$ of dimension 2; any irreducible curve on the surface corresponds to a prime ideal $\mathfrak{p}_1$ dividing $\mathfrak{p}_2$, while every point on the curve is a zero-dimensional prime ideal $\mathfrak{p}_0$ dividing $\mathfrak{p}_2$ and $\mathfrak{p}_1$.

Since the manifolds defined by prime ideals are known, the manifolds of all other ideals can be determined; the manifold of a primary ideal $\mathfrak{q}$ is equal to the manifold of its prime ideal $\mathfrak{p}$, since $\mathfrak{p}$ divides $\mathfrak{q}$ and a power of $\mathfrak{p}$ is divisible by $\mathfrak{q}$; the manifold of the union of two ideals is equal to the union of the manifolds. If we represent an ideal $\mathfrak{a}$ as the union of maximal primary components, *the manifold of* $\mathfrak{a}$ *must then be equal to the union of the manifolds defined by the prime ideals belonging to* $\mathfrak{a}$. The fact that these prime ideals were uniquely defined is only a new way of stating that every algebraic manifold can be represented uniquely as the union of irreducible manifolds.

These remarks give us a simple access to various fundamental theorems in the theory of algebraic manifolds. Let $f(x)$ be a polynomial vanishing at all the points of the manifold $\mathfrak{M}$ defined by the relations (12); $f(x)$ is then divisible by all prime ideals belonging to the ideal

$$\mathfrak{a} = (f_1(x), \cdots, f_r(x)), \tag{13}$$

and if ρ is an upper bound for the exponents of the primary components of $\mathfrak{a}$, then $f(x)^\rho$ must belong to $\mathfrak{a}$. This is equivalent to a theorem of Hilbert: *If a polynomial* $f(x_1, \cdots, x_n)$ *vanishes*

at all points satisfying (12), *then there exists an exponent such that*

$$f(x)^\rho = g_1(x)f_1(x) + g_2(x)f_2(x) + \cdots + g_r(x)f_r(x).$$

Another important problem is to determine when a polynomial $f(x)$ belongs to a given ideal $\mathfrak{a}$; we first consider the simplest case of an ideal $\mathfrak{a}=(f_1(x, y), f_2(x, y))$ in two variables and suppose that the curves $f_1(x, y)=0$, $f_2(x, y)=0$ have only a finite number of points $(x_1, y_1), \cdots, (x_r, y_r)$ in common. The manifold of $\mathfrak{a}$ is then determined by the zero-dimensional prime ideals $\mathfrak{p}=(x-x_i, y-y_i)$, and $\mathfrak{a}$ has a representation of the form $\mathfrak{a}=[\mathfrak{g}_1, \cdots, \mathfrak{g}_r]$ as the union of maximal primary components $\mathfrak{g}_i$ with the corresponding prime ideals $\mathfrak{p}$. For a polynomial $f(x, y)$ to belong to $\mathfrak{a}$ it is necessary and sufficient that it belong to all primary components $\mathfrak{g}_i$; the ideals $\mathfrak{g}_i$ are not known directly, but since $\mathfrak{g}$ divides $\mathfrak{p}_i^{\rho_i}$, one finds that $\mathfrak{g}_i=(\mathfrak{a}, \mathfrak{p}_i^{\rho_i})$ and we obtain the condition of M. Noether,

$$f(x, y) \equiv 0 \pmod{(\mathfrak{a}, \mathfrak{p}_i^{\rho_i})}. \tag{14}$$

Since $\mathfrak{p}_i^{\rho_i}$ consists of all terms of the form $\Sigma c_{a,b}(x-x_i)^a(y-y_i)^b$, $a+b \geqq \rho_i$, it follows that $f(x, y)$ belongs to $\mathfrak{a}$, that is, is representable in the form

$$f(x, y) = g_1(x, y)f_1(x, y) + g_2(x, y)f_2(x, y) \tag{15}$$

if and only if such a representation (15) exists for all intersections when terms of degree ρ in $x-x_i$ and $y-y_i$ are disregarded. The importance of the theorem of M. Noether for algebraic geometry is well known; a condition similar to (14) for an arbitrary ideal (13) was given by Henzelt.*

10. *Unique Prime Ideal Decomposition.* It is a problem of considerable interest to determine the necessary and sufficient conditions which a ring must satisfy in order that every ideal be representable uniquely as the product of a finite number of prime ideal factors. This problem has been investigated systematically by E. Noether.† We shall consider a ring R with

* Henzelt, Mathematische Annalen, vol. 88 (1923); E. Noether, Mathematische Annalen, vol. 90 (1923); Hermann, Mathematische Annalen, vol. 95 (1925); Kapferer, Mathematische Annalen, vol. 97 (1927); v. d. Waerden, Mathematische Annalen, vol. 96 (1927) and vol. 99 (1928), *Algebra*, Chap. 13.

† Mathematische Annalen, vol. 96 (1927).

unit element and without divisors of zero. It is obvious that it is necessary first to impose the *chain condition* on the ring; secondly it must be assumed that the *prime ideals in R have no divisors* $\neq R$. This condition either may be postulated or may be derived as a consequence of the so-called *chain condition for multipla of ideals. Every chain* $\mathfrak{a}_1 > \mathfrak{a}_2 > \cdots$ *of ideals, each divisible by the preceding, must break off when all the ideals in the chain are divisors of a fixed ideal* $\mathfrak{a} \neq 0$. When the two chain conditions are satisfied the three decompositions for ideals given in §8 are identical and every ideal in R is representable uniquely as the product of its maximal primary components. It remains then to find the condition which will make every primary ideal equal to a prime ideal power; this condition is closely connected with the notion of *integral elements.*

Let S and T be two rings; S shall be a subring of T and have a unit element. An element t in T is said to be *integral with respect to* S, when the powers of t can be expressed by a basis $t_1, \cdots, t_n$ in the form

$$t^m = s_1^{(m)} t_1 + \cdots + s_n^{(m)} t_n, \tag{16}$$

with coefficients $s_j^{(i)}$ in S. When S satisfies the chain condition, (16) is equivalent to the condition that t shall satisfy an algebraic equation

$$t^n + s_1 t^{n-1} + \cdots + s_n = 0.$$

We say that S is *integrally closed* with respect to T when every element in T which is integral with respect to S is already contained in S.

Let us now return to the ring R considered above, and impose the third condition: *R shall be integrally closed with respect to its quotient field.* Every primary ideal can then be shown to be a prime ideal power and there is a unique prime ideal decomposition of each ideal in R. The theory is completed by the converse theorem, that if there exists a unique prime ideal decomposition then the ring must satisfy the three conditions mentioned.*

Some of the most important rings in algebra are integrally closed and satisfy the chain condition, but prime ideal divisors

* See also Krull, Sitzungsberichte Heidelberg, 1924, 1925; Mathematische Annalen, vol. 99 (1928).

may exist. The rings of polynomials are of this type. In this case, however, a certain class of ideals, namely the principal ideals, have a unique decomposition into prime factors. Van der Waerden* has pointed out that the decomposition theorem holds even for rings with prime ideal divisors, when only the so-called *highest prime ideals* are considered. A highest prime ideal is then defined as a prime ideal not dividing any other prime ideal; in the case of polynomial ideals these prime ideals define the algebraic manifolds of dimension $n-1$ and are ordinary prime polynomials. Through a suitable definition of equivalence of ideals van der Waerden shows that every primary ideal is equivalent to a prime ideal power, and that every ideal is equivalent to a unique product of prime ideals; in the case of rings without prime ideal divisors this reduces to the ordinary decomposition theorem.

11. *Rings without Chain Condition.* We shall conclude this survey by considering briefly the results obtained for the ideal theory in rings without chain condition. The examples which lie closest at hand are the infinite algebraic fields; the ideal theory of the ring of all integers in such fields has been discussed in a remarkable posthumous paper by Stiemke.† There exist prime ideals and also an enumerable basis for each ideal. One can usually not expect any representation of ideals as product of prime ideals since the prime ideal divisors of even a single prime may form a set having the power of the continuum. There are, however, special infinite algebraic fields in which there exists only an enumerable set of prime ideals and in such fields every integral element can be represented uniquely as the product of prime ideal factors and certain primary factors, which can be characterized by fractional exponents. An example of a field satisfying the condition of Stiemke is the field of all p^mth roots of unity ($m=0, 1, 2, \cdots$), where p is a fixed prime.

The investigations of Stiemke have been continued by Krull,‡ who has obtained a method of handling the set of ideals in an infinite ring of algebraic numbers through the introduction

* Mathematische Annalen, vol. 99 (1928); *Algebra*, §103.

† Mathematische Zeitschrift, vol. 25 (1926).

‡ Mathematische Zeitschrift, vol. 29 (1929); vol. 31 (1930).

and systematic use of the notion of *absolute values.** Let K be any field. We say that an absolute value is defined in K when to each element α in K there corresponds a real number $\|\alpha\|$ with the properties

$$\|0\| = 0,\ \|\alpha\| > 0,\ \text{when } \alpha \neq 0, \qquad \|\alpha\beta\| = \|\alpha\| \cdot \|\beta\|,\ \|\alpha + \beta\| \leqq \|\alpha\| + \|\beta\|. \tag{17}$$

Let us consider the rational field P as an example. In this field there are only two different types of absolute values; the first is the trivial type

$$\|\alpha\| = |\alpha|^{\rho},\ 0 \leqq \rho \leqq 1, \tag{18}$$

where ρ is a constant and $|\alpha|$ denotes the ordinary absolute value. To obtain the second type let p be a prime and let us write all rational numbers in the form

$$r = p^{a}\frac{u}{v},\ (u, v) = 1, \tag{19}$$

where u and v are integers not divisible by p; the correspondence

$$\|r\| = c^{a},\ 0 < c \leqq 1, \tag{20}$$

then satisfies the conditions for an absolute value. The main difference between the two types (18) and (20) of absolute values lies in the *triangular inequality* (17); in the case of the type (20) the inequality may be strengthened to

$$\|\alpha + \beta\| \leqq \text{Max}\,(\|\alpha\|, \|\beta\|) \tag{21}$$

and we say that the absolute value is *non-Archimedian*. Ostrowski† has shown that every field with Archimedian absolute values is isomorphic to a subfield of the field of all complex numbers, and that the absolute value can then be defined as in (18).

For the study of divisibility properties only the non-Archimedian absolute values are, therefore, of importance; in this case the exponent a in (19) is the characteristic to be investi-

* See for instance the exposition of this theory in O. Ore, *Some recent developments in abstract algebra*, this Bulletin, vol. 37 (1931), pp. 537–548.

† Acta Mathematica, vol. 41 (1917).

gated and it is, consequently, more convenient to define the absolute value $v(\alpha)$ in the *additive form*

$$(22) \qquad v(\alpha\beta) = v(\alpha) + v(\beta),\ v(\alpha + \beta) \geqq \text{Min}\,(v(\alpha), v(\beta)).$$

One obtains (22) from (17) and (21) by putting $v(\alpha) = -\log\|\alpha\|$, and for the rational case (19) it reduces to $v(r) = ka$, where the positive constant k may be taken equal to unity.

We return again to the representation (19) of the rational numbers; the numbers r for which the exponent a is non-negative form a ring P_p which we shall call an *evaluation ring* (*Bewertungsring*) with respect to p. This ring has a series of properties which all have their analogues in the more general evaluation rings considered later; P_p is a maximal ring in P, that is, there is no ring containing P_p except P; P_p is integrally closed in P; when r_1 and r_2 are two elements in P_p, then r_1 is divisible by r_2, or conversely; the ideals in P_p are all of the form $\mathfrak{a} = (p^n)$ and an absolute value for the ideals can be defined by putting $v(\mathfrak{a}) = n$; the ideal $\mathfrak{a}$ consists of all elements r such that $v(r) \geqq v(\mathfrak{a})$; the only prime ideal is $\mathfrak{p} = (p)$ consisting of all elements with positive absolute values. The ring of rational integers is equal to the cross-cut of all rings P_p for all p, and the divisibility properties of any rational integer are determined by its values in the various evaluation rings.

Krull now shows that for each definition of absolute value in a field there exists a corresponding evaluation ring consisting of the elements with non-negative values; this ring is maximal and conversely every maximal ring defines an absolute value. These rings have the properties mentioned above: An evaluation ring is integrally closed; when two elements are given, one always divides the other; there exists only one prime ideal $\mathfrak{p}$ consisting of all elements having positive absolute values. There are now two different possibilities. The values of the elements may form a discrete set; in this case we have perfect analogy to the rational case and all ideals are powers of $\mathfrak{p}$. In the second case the values may be everywhere dense; in both cases the ideals $\mathfrak{a}$ consist of the elements whose values lie above a certain limit κ. We define a value of $\mathfrak{a}$ by putting $v(\mathfrak{a}) = \kappa$; to each κ there may correspond two ideals, namely, the ideal $\mathfrak{a}$ containing all α such that $v(\alpha) > v(\mathfrak{a}) = \kappa$ and the ideal $\mathfrak{b}$ containing also elements β such that $v(\beta) = v(\mathfrak{b}) = \kappa$. We shall call

$\mathfrak{a}$ an *open ideal* and $\mathfrak{b}$ a *closed ideal* and say that $v(\mathfrak{a}) > v(\mathfrak{b})$; in the case of continuous values $\mathfrak{p}$ is an open ideal with the value 0 and the unit ideal is a closed ideal with the same value.

This theory may be applied to the ring R of all integers in an infinite algebraic field K. One can show that the maximal rings in K are defined by the prime ideals $\mathfrak{p}$ in R and that $R_\mathfrak{p}$ consists of all fractional elements in K whose denominators are not divisible by $\mathfrak{p}$; and that the ring of integers R is equal to the cross-cut of all maximal rings $R_\mathfrak{p}$ as in the rational field. To each ideal $\mathfrak{a}$ in R corresponds an ideal $\mathfrak{a}_\mathfrak{p}$ in $R_\mathfrak{p}$ and this permits us to define a set of values $v_\mathfrak{p}(\mathfrak{a})$ for each ideal $\mathfrak{a}$, where $\mathfrak{p}$ runs through all prime ideals of R. These values may be considered as a *generalization of the exponents* in the case of prime ideal decompositions in the finite fields. They satisfy the addition theorem, $v_\mathfrak{p}(\mathfrak{ab}) = v_\mathfrak{p}(\mathfrak{a}) + v_\mathfrak{p}(\mathfrak{b})$, and the fundamental theorem is, that *an ideal* $\mathfrak{a}$ *can only be divisible by an ideal* $\mathfrak{c}$ *when* $v_\mathfrak{p}(\mathfrak{a}) \geqq v_\mathfrak{p}(\mathfrak{c})$ *for all prime ideals* $\mathfrak{p}$; one can also find a necessary and sufficient condition that an ideal be a factor of another. The inverse problem, namely, finding all sets of values to which there correspond ideals, Krull was able to solve, through topological considerations, by making the set of all evaluation rings a topological space.

In a recent paper, Krull* generalizes the notion of absolute value by letting correspond to each element of a field not a real number as formerly, but an element of an ordered abelian group, such that the conditions (22) are satisfied. This absolute value corresponds to a generalized type of evaluation rings with a number of properties in common with the evaluation rings studied above. One of the most interesting results of this theory is the following characterization of the integrally closed rings. *A necessary and sufficient condition that a ring be integrally closed is that it be the cross-cut of general evaluation rings.* Krull also gives various other results which cannot be discussed here; a paper by Prüfer† on the divisibility of ideals in rings without chain condition and a paper by v. Dantzig‡ on topological algebra should be mentioned in this connection.

YALE UNIVERSITY

* Journal für Mathematik, vol. 167 (1932).

† Journal für Mathematik, vol. 168 (1932).

‡ Mathematische Annalen, vol. 107 (1932).

NORMAL DIVISION ALGEBRAS OVER ALGEBRAIC NUMBER FIELDS NOT OF FINITE DEGREE*

BY A. A. ALBERT

1. *Introduction.* If R is the field of all rational numbers and if $\xi_1, \cdots, \xi_n$ are ordinary algebraic numbers, then the field $\Omega = R(\xi_1, \cdots, \xi_n)$ of all rational functions with rational coefficients of $\xi_1, \cdots, \xi_n$ is an *algebraic number field of finite degree* (the maximum number of linearly independent quantities of Ω) *over* R. It has recently been proved† that every normal *simple* algebra over such a field Ω is cyclic. In particular it has been shown that every normal *division* algebra of order n^2 (degree n) over Ω is cyclic and has exponent n.

In the present note I shall give an extension of the above results to *normal division algebras over any algebraic number field* Λ. I shall prove that all normal division algebras over Λ are cyclic and with degree equal to exponent but shall give a trivial example showing that the theorem corresponding to the above on normal *simple* algebras is false. The problem of the equivalence of normal division algebras over Λ will also be discussed.

2. *Cyclic Algebras.* Let F be any non-modular field and let Z by cyclic of degree n over F. Then Z possesses a generating automorphism

$$S: \qquad z \longleftrightarrow z^S, \qquad (z \text{ in } Z,\ z^S \text{ in } Z),$$

such that every automorphism of Z is one of $S^0 = S^n = I$, S, $S^2, \cdots, S^{n-1}$. The algebra A of all quantities

$$\sum_{i=0}^{n-1} z_i y^i, \qquad (z \text{ in } Z),$$

is a cyclic algebra with multiplication table

$$y^n = \gamma \text{ in } F,\ y^e z = z^{S^e} y^e, \qquad (e = 0, 1, \cdots),$$

* Presented to the Society, October 28, 1933.

† See the paper by H. Hasse and myself in the Transactions of this Society, vol. 34 (1932), pp. 722–726, for the normal division algebra theorem. The theorem for normal simple algebras follows from Hasse's Theorem 6 of his Transactions paper, vol. 34 (1932), pp. 171–214.

for every z of Z. Evidently A is uniquely defined by Z, S, γ, and thus we write

$$A = (Z, S, \gamma).$$

Let F be contained in any larger field K. Then

$$A_K = (Z, S, \gamma)_K$$

is the algebra with the same basis and constants of multiplication as A, but over K.

If A_K is a division algebra, then so evidently is A. But then Z_K, which is the algebra with the same basis and constants of multiplication as the field Z, but over K, is a field and in fact is evidently cyclic of degree n over K. Evidently $A = (Z_K, S, \gamma)$ over K.

THEOREM 1. *Let $A = (Z, S, \gamma)$ over F, $F < K$, and let A_K be a division algebra. Then A_K is the cyclic algebra (Z_K, S, γ) over K.*

3. *The Determination of Algebras over* Λ. Let Λ be any non-modular field whose quantities are all algebraic numbers and let A be a normal division algebra of order $m = n^2$ over Λ. If $u_1, \cdots, u_m$ are a basis of A, then $u_i u_j = \sum' \gamma_{ijk} u_k$ with γ_{ijk} in Λ. But then γ_{ijk} are all algebraic numbers, so that $L = R(\gamma_{111}, \cdots, \gamma_{ijk}, \cdots, \gamma_{mmm})$ is algebraic of finite degree.

The linear set $B = (u_1, \cdots, u_m)$ over L is evidently an algebra of order m over L. If in particular $u_1 = 1$, the modulus of A, then u_1 is the modulus of B. Evidently $A = B_\Lambda$.

If B contains any divisors of zero, then these quantities are in the division algebra A, a contradiction. Hence B is a division algebra.

Let B contain a quantity $k = \sum \lambda_i u_i$, λ_i in L, which is commutative with every quantity of B. In particular $ku_i = u_i k$, so that $k(\sum \mu_i u_i) = (\sum \mu_i u_i)k$ for μ_i any quantities of the field Λ. But A is normal, so that k is a multiple of the modulus u_1 of A by a quantity of Λ. Hence $k = \mu u_1 = \sum \lambda_i u_i$. Since the u_i are linearly independent in Λ, we have $\mu = \lambda_1$, k is a multiple of u_1 by a quantity of L, and B is normal.

The normal division algebra B of degree n over L is thus* a cyclic algebra (Z, S, γ) over L. The basis, (u_i), of A is linearly

* By the result already quoted on normal division algebras over Ω.

expressible with coefficients in L in terms of the basis of $B=(Z, S, \gamma)$ in its cyclic form, so that in fact $A=(Z, S, \gamma)_\Lambda$. By Theorem 1 we have the following result.

THEOREM 2. *Let A be a normal division algebra of degree n over an algebraic number field Λ not of finite degree. Then there exists a sub-field L (of Λ) of finite degree and a cyclic algebra, $B=(Z, S, \gamma)$, over L such that $A=(Z_\Lambda, S, \gamma)$ over Λ, where Z_Λ is a cyclic field of degree n over Λ. Hence A is cyclic.*

4. *The Exponent of Algebras A.* Suppose that the algebra A of Theorem 2 has exponent $\rho<n$. Then A^ρ is well known to be equal to $M^{\rho-1}\times(Z_\Lambda, S, \gamma^\rho)$, where M is a total matric algebra. But A^ρ is a total matric algebra; hence $(Z_\Lambda, S, \gamma^\rho)$ is also. Hence γ^ρ is the norm $N(c)$ of a quantity c of Z_Λ.

Let $Z=L(x)$, $Z_\Lambda=\Lambda(x)$, so that $c=\sum c_i x^i$, where the c_i are in Λ. The field $L=L(c_0, \cdots, c_{n-1})$ is algebraic of finite degree. Moreover, if $B=(Z, S, \gamma)$, then evidently $Z_0=L_0(x)$, $B_0=(Z_0, S, \gamma)$ over L_0, is contained in A and hence is a cyclic division algebra. But $B_0^\rho=(Z_0, S, \gamma^\rho)\times M^{\rho-1}$ is a total matric algebra, since $\gamma^\rho=N(c)$, where c is in Z_0.

The exponent of B_0 of degree n over L_0 is known to be n since B is a cyclic division algebra over L_0, which is *algebraic of finite degree*. Hence $\rho\geqq n$, a contradiction.

THEOREM 3. *The exponent of any normal division algebra over Λ is its degree.*

5. *On the Equivalence of Algebras over Λ.* Let $A=(Z_\Lambda, S, \gamma)$ and $C=(Y_\Lambda, T, \delta)$ over Λ be normal division algebras. Then Z and γ are obtained with respect to a field L_1 defined by A, Y, and δ with respect to L_2 defined by C. If L is the composite of L_1 and L_2, then we may evidently take L as the common field of Theorem 2 for both algebras A and C. Hence $A=(Z, S, \gamma)_\Lambda$, (Z, S, γ) a normal division algebra over L, $C=(Y, T, \delta)_\Lambda$, (Y, T, δ) also a normal division algebra over L.

The algebra A is equivalent to the algebra C if and only if $A\times C^{-1}=(Z, S, \gamma)\times(Y, T, \delta^{-1})$ is a total matric algebra. But, as is well known, $(Z, S, \gamma)\times(Y, T, \delta^{-1})=(X, R, \epsilon)\times M$, where M is a total matric algebra and (X, R, ϵ) is a uniquely determined cyclic algebra. Evidently $A\times C^{-1}$ is total matric if and only if $(X, R, \epsilon)_\Lambda$ is total matric. For $A\times C^{-1}=M\times(X, R, \epsilon)_\Lambda$.

But then $\epsilon = N(c)$, where c is in X_Λ. As before there exists a sub-field L_0 of Λ of finite degree such that c is in X_{L_0}, $(X, R, \epsilon)_{L_0}$ is total matric. But then $(Z, S, \gamma)_{L_0}$ is equivalent to $(Y, T, \delta)_{L_0}$. The converse is obvious and we have proved this theorem.

THEOREM 4. *Let A and C be normal division algebras of degree n over Λ, an algebraic field not of finite degree, so that $A = (Z_\Lambda, S, \gamma)$, $C = (Y_\Lambda, T, \delta)$, where $B = (Z, S, \gamma)$, $D = (Y, T, \delta)$ are cyclic over the same sub-field L of finite degree of Λ. Then A and C are equivalent if and only if there exists a sub-field $L_0 > L$ of Λ such that L_0 has finite degree and the algebras B_{L_0} and D_{L_0} are equivalent.*

The above theorem essentially reduces the problem of the equivalence of normal division algebras over A to the corresponding problem (already solved*) for algebras over fields of finite degree, and to a consideration of the sub-fields of Λ of finite degree.

6. *Normal Simple Algebras over* Λ. In this section we shall show trivially that there exist non-cyclic normal simple algebras over an algebraic field Λ. We take Λ to be the field of all constructible (with ruler and compass) numbers, extended by $i = (-1)^{1/2}$. That is, we take Λ to consist of all numbers obtained from rational numbers by a finite number of rational operations and extractions of square roots.

Evidently any equation $x^2 = c$, c in Λ, is reducible in Λ since $c^{1/2}$ is also in Λ. But then there exist no cyclic algebras of degree *two* over Λ. Hence the total matric algebra of degree two over Λ, a normal simple algebra, is non-cyclic.

THE INSTITUTE FOR ADVANCED STUDY

* See Hasse, loc. cit.

FURTHER MEAN-VALUE THEOREMS*

BY MORRIS MARDEN

The present note is a sequel to my recent article† in which certain mean-value theorems due to Weierstrass and Fekete were generalized. The generalizations resulted from replacing a positive real weight-function by one assuming only values in the angular region $0 \leqq \arg w \leqq \gamma < \pi$. Here the generalizations will be extended in such a manner as to yield analogous theorems in which the weight-function takes on arbitrary real values (Corollary 2, Theorem 3) or more generally any values in the double angle $0 \leqq \arg (\pm w) \leqq \gamma < \pi$ (Corollary 1, Theorem 3). Incidentally, these extensions yield (as Corollary 3, Theorem 3) the generalization of the Gauss-Lucas theorem which formed the principal result of another paper.‡

In what follows we shall denote by $f(Z)$ the point set $w = f(z)$ obtained on letting point z vary over the point set Z; by $\Delta \arg (Z - p)$ the magnitude of the smallest angle, with vertex at the point p, enclosing the point set Z; by $K(Z)$ the smallest convex region containing set Z, and, finally, by $S(Z, \theta)$ the star-shaped region composed of all points from which the set Z subtends an angle of not less than θ. The regions $K(Z)$ and $S(Z, \theta)$ can also be defined as the loci of all points p which satisfy respectively the inequalities $\Delta \arg (Z - p) \geqq \pi$, $\Delta \arg (Z - p) \geqq \theta$. Obviously, $S(Z, \theta) \equiv S(K(Z), \theta)$ and hence $S(Z, \theta)$ always contains $K(Z)$. Finally, in what follows, the two rectifiable curves

$$C: \quad z = z(s), \quad a \leqq s \leqq b; \qquad \Gamma: \quad \lambda = \lambda(t), \quad \alpha \leqq t \leqq \beta,$$

will serve as the curves of integration, and, unless further qualified, all functions introduced hereafter will be supposed to be continuous on these curves except perhaps for a finite number of finite jumps.

---

* Presented to the Society, April 15, 1933.

† M. Marden, this Bulletin, vol. 38 (1932), pp. 434–441.

‡ M. Marden, *On the zeros of certain rational functions*, Transactions of this Society, vol. 32 (1930), pp. 658–668.

THEOREM 1. *Let there be given the real numbers m_i and the functions $f_i(z)$ and $g(z)$ with $\Delta \arg g(C) \leqq \gamma < \pi$. Then each point σ as defined by the equation*

$$\int_a^b g(z) \prod_{i=1}^{h} [f_i(z) - \sigma]^{m_i} ds = 0 \tag{1}$$

lies in the region $S[K(f_1(C), f_2(C), \cdots, f_h(C)), (\pi-\gamma)/m]$, where $m = \sum_1^h |m_i|$.

THEOREM 2. *Let, in particular, each $f_i(z)$ be a rational function with exactly n_i finite zeros and p_i finite poles, none of the latter lying in the region $S(C, (\pi-\gamma)/n)$, where $n = \sum_1^h |m_i| (p_i+q_i)$ and $q_i = \max(n_i, p_i)$. Then for each value σ as defined by (1) there exists at least one integer i, $1 \leqq i \leqq h$, and at least one point z in $S(C, (\pi-\gamma)/n)$, such that $f_i(z) = \sigma$.*

Suppose Theorem 1 were not true; that is, suppose that, for some σ and for all i,

$$\Delta \arg [f_i(C) - \sigma] < \frac{\pi - \gamma}{m}.$$

Then

$$\Delta \arg [f_i(C) - \sigma]^{m_i} < \frac{\pi - \gamma}{m} |m_i|,$$

and hence

$$\Delta \arg g(C) \prod_{i=1}^{h} [f_i(C) - \sigma]^{m_i} < \pi.$$

Accordingly, the left hand-side of (1) is a sum of vectors each drawn from $w = 0$ to points on the same side of some line through $w = 0$. As such a sum cannot vanish, the assumption that Theorem 1 is false contradicts equation (1). Hence Theorem 1 must be true.

Similarly, let us suppose Theorem 2 to be false; that is, writing

$$f_i(z) - \sigma = A_i \frac{(z - a_{i1})(z - a_{i2}) \cdots (z - a_{iq_i})}{(z - b_{i1})(z - b_{i2}) \cdots (z - b_{ip_i})},$$

let us suppose that for all j and k

$$\Delta \arg (C - a_{jk}) < \frac{\pi - \gamma}{n}.$$

Since we know by hypothesis, for all j and k, that

$$\Delta \arg (C - b_{jk}) < \frac{\pi - \gamma}{n},$$

it would follow that

$$\Delta \arg [f_i(C) - \sigma] < \frac{\pi - \gamma}{n}(p_i + q_i),$$

and hence

$$\Delta \arg g(C) \prod_{i=1}^{k} [f_i(C) - \sigma]^{m_i} < \pi.$$

Again equation (1) would be contradicted and hence Theorem 2 must be true.

On setting each $m_i = h = 1$, we obtain from Theorem 1 a previous generalization* of Weierstrass' mean-value theorem, and on setting also $\gamma = 0$, we derive his original theorem.†

The choice $m_i - 1 = p_i = h - 1 = 0$ for all i reduces Theorem 2 to the previous generalization of Fekete's theorems,‡ particularly of his following two theorems.

(1) *If $f(z)$ is a polynomial of degree n and $f(\alpha) \neq f(\beta)$, $\alpha \neq \beta$, it assumes every value between $f(\alpha)$ and $f(\beta)$, that is, on the line-segment joining $f(\alpha)$ and $f(\beta)$, at least once in S* (seg $\alpha\beta$, π/n).

(2) *If $P(z)$ is a polynomial of degree n and $P(\alpha) = P(\beta)$, $\alpha \neq \beta$, then $P'(z) = 0$ at least once in S*(seg $\alpha\beta$, $\pi/(n-1)$).

The first of Fekete's theorems is analogous to the Bolzano theorem for continuous functions of a real variable. The second Fekete theorem is analogous to Grace's theorem§, that under the same assumptions, $P'(z) = 0$ at least once in the circle with its center at $(\alpha+\beta)/2$ and with a radius of $\frac{1}{2}|\beta - \alpha|$ ctn (π/n). It is interesting to note that, since the circle of Grace's theorem passes through the centers of the two circles bounding

* M. Marden, this Bulletin, loc. cit., p. 435.

† Osgood, *Lehrbuch der Funktionentheorie*, 1923, vol. I, p. 212.

‡ See this Bulletin, loc. cit., p. 438 and p. 440. Also M. Fekete, Acta Szeged, vol. 1 (1923), pp. 98–100, and vol. 4 (1929), pp. 234–243; Mathematische Zeitschrift, vol. 22 (1925) pp. 1–7; Jahresbericht der Deutschen Mathematiker-Vereinigung, vol. 32 (1923), pp. 299–306, and vol. 34 (1926), pp. 220–233. J. v. Sz. Nagy, Jahresbericht der Deutschen Mathematiker-Vereinigung, vol. 32 (1923), pp. 307–309.

§ P. J. Heawood, Quarterly Journal of Mathematics, vol. 38 (1907), pp. 84–107.

$S(\text{seg } \alpha\beta, \pi/(n-1))$, a better approximation to a zero of $P'(z)$ is obtained through use of both Grace's and Fekete's theorems than through either separately.

Theorem 1 may be stated in the following more general form.

THEOREM 3. *Let there be given the functions $f_i(z, \lambda)$ and $g(z, \lambda)$ with $\Delta \arg g(C, \Gamma) \leqq \gamma < \pi$. Then each point σ defined by the equation*

$$\int_\alpha^\beta \int_a^b g(z, \lambda) \prod_{i=1}^{h} [f_i(z, \lambda) - \sigma]^{m_i} ds dt = 0, \tag{2}$$

lies in

$$S\left[K(f_1(C, \Gamma), f_2(C, \Gamma), \cdots, f_h(C, \Gamma)), \frac{\pi - \gamma}{m}\right],$$

where $m = \sum_1^h |m_i|$.

This theorem may be proved precisely as was Theorem 1.

If in Theorem 3 we specialize $\lambda(t) \equiv 0$ for $\alpha = 0 \leqq t \leqq 1 = \beta$ and $\lambda(t) \equiv 1$ for $1 < t \leqq 2 = \beta$ and $f(z, 0) = -f(z, 1) = f(z)$, and if we let $g_1(z) = g(z, 0)$ and $g_2(z) = g(z, 1)$, we derive the following result.

COROLLARY 1. *Let there be given the functions $f(z)$, $g_1(z)$, and $g_2(z)$ with $0 \leqq \arg g_i(z) \leqq \gamma < \pi$ for $i = 1, 2$; then the point σ, as defined by the equation*

$$\int_a^b [g_1(z) - g_2(z)] f(z) ds = \sigma \int_a^b [g_1(z) + g_2(z)] ds,$$

lies in

$$S(\pm f(C), \pi - \gamma).$$

This result leads us to a mean-value theorem in which the weight-function $g(z)$ is real, but not necessarily positive. We may indeed define two functions $g_1(z)$ and $g_2(z)$ so that

$$g_1(z) \equiv g(z), \quad g_2(z) \equiv 0, \qquad \text{for} \quad g(z) \geqq 0,$$
$$g_1(z) \equiv 0, \quad g_2(z) \equiv -g(z), \quad \text{for} \quad g(z) \leqq 0.$$

These functions $g_1(z)$ and $g_2(z)$ fulfill the requirements of Corollary 1 with $\gamma = 0$ and

$$g_1(z) - g_2(z) = g(z), \qquad g_1(z) + g_2(z) = |g(z)|.$$

The resulting theorem may be stated as follows.

COROLLARY 2. *Let there be given the complex function $f(z)$ and the real function $g(z)$. Then the point σ, as defined by the equation*

$$\int_a^b g(z)f(z)ds = \sigma \int_a^b |\, g(z)\,|\, ds,$$

lies in $K(\pm f(C))$.

Finally in Theorem 3 let us specialize as follows:

$$\alpha = 0, \beta = 1, a = 0, b = r \text{ (an integer)};$$
$$z(s) \equiv z_j \text{ for } j - 1 \leqq s < j;$$
$$\lambda(t) \equiv 0; \qquad g(z_j, 0) = \alpha_j;$$
$$m_k = 1, \quad f_k(z_j, 0) = a_{jk} \text{ for } \quad 1 \leqq k \leqq n;$$
$$m_k = -1, \; f_k(z_j, 0) = b_{jk} \text{ for } \; n + 1 \leqq k \leqq n + m = h;$$

and thus obtain the following corollary.*

COROLLARY 3. *Let α_i be any complex numbers such that for all i, $0 \leqq \arg \alpha_i \leqq \gamma < \pi$, and let a_{jk} and b_{jk} for all j and k be points of a given convex region K. Then all the zeros of the function*

$$\Phi(z) = \sum_1^r \alpha_i \phi_i(z),$$

where

$$\phi_i(z) = \frac{(z - a_{i1})(z - a_{i2}) \cdots (z - a_{in})}{(z - b_{i1})(z - b_{i2}) \cdots (z - b_{im})},$$

lie in the region

$$S\left(K, \frac{\pi - \gamma}{m + n}\right).$$

The particular case of this corollary $\gamma = \alpha_i - 1 = n = m - 1 = 0$ yields the theorem that the zeros of the partial fraction sum $\sum_1^r (z - z_i)^{-1}$ lie in the smallest convex region enclosing the points z_i. As this partial fraction is the logarithmic derivative of the polynomial $f(z) = A(z - z_1)(z - z_2) \cdots (z - z_r)$, this special case is identical with the Gauss-Lucas theorem for the zeros of the derivative of a polynomial.

UNIVERSITY OF WISCONSIN,
MILWAUKEE EXTENSION CENTER

* See Marden, Transactions of this Society, loc. cit.

ON THE NUMBER OF APPARENT DOUBLE POINTS ON A CERTAIN V_k^n IN AN S_{2k+1}

BY B. C. WONG

Consider a k-dimensional variety, V_k^n, of order n which is the locus of a single infinity of $(k-1)$-spaces in an S_{2k+1}. It is known that such a V_k^n, if it is rational, that is, if its section by a general S_{k+2} of S_{2k+1} is a rational curve, has*

$$b_k = \tfrac{1}{2}(n-k)(n-k-1)$$

apparent double points.† The question arises: What is the value of b_k when V_k^n is not rational? The case $k=1$ is familiar; a curve of order n in an S_3 has

$$b_1 = \tfrac{1}{2}(n-1)(n-2) - p \tag{1}$$

apparent double points, where p is the deficiency of the curve. It is also known that, for $k=2$, the number of apparent double points on a ruled surface F^n of order n in an S_5 is‡

$$b_2 = \tfrac{1}{2}(n-2)(n-3) - 3p, \tag{2}$$

where p is the deficiency of the curve of intersection of F^n by a general S_4 of S_5. For $k>2$, the number b_k of apparent double points of a V_k^n in an S_{2k+1} seems to be as yet unknown. It is our purpose in this note to derive a formula for this number.

Now let V_k^n be intersected by a general S_{k+2} of S_{2k+1} in a curve C^n of deficiency p. If $p>0$, we say that V_k^n is not rational. We shall say that p is also the deficiency of V_k^n and shall regard n and p as the two essential characteristics of the variety as all its other characteristics can be expressed in terms of them for a

* B. C. Wong, *On the number of $(q+1)$-secant S_{q-1}'s of a certain V_k^n in an $S_{qk+q+k+1}$*, this Bulletin, vol. 39, pp. 392–394.

† By an apparent double point of a V_k^n we mean a secant line of V_k^n passing through a given point of S_{2k+1}. The projection in an S_{2k} of V_k^n will have b_k improper double points each of which is the projection of an apparent double point of V_k^n.

‡ Severi, *Intorno ai punti doppi impropri di una superficie generale dello spazio a quattro dimensioni, e a'suoi punti tripli apparenti*, Rendiconti di Palermo, vol. 15 (1901), pp. 33–51.

given value of k. Then the formula for b_k must be a function of n and p, and of k also.

Consider the ruled surface F^n in which V_k^n is met by a general S_{k+2} of S_{2k+1}. The projection of F^n, if it is in an S_5, has b_2, given by formula (2), apparent double points; and, if it is in an S_3, has a double curve of order b_1 given by formula (1). On this double curve lie a finite number, j_1, of pinch points. This number is known and will be given subsequently.

Next, consider the planed variety V_3^n common to V_k^n and a general S_{k+4} of S_{2k+1}. The projection in an S_7 of this V_3^n has b_3 apparent double points. Projecting this projection successively upon an S_6, an S_5, and an S_4, we see that the resulting variety in S_6 has b_3 improper double points; that the one in S_5 contains a double curve of order b_2 upon which lie j_2 pinch points; and, finally, that the one in S_4 contains a double surface of order b_1 upon which lies a pinch curve of order j_1.

In general, an $S_{k+h+1}(k \geqq h > 0)$ of S_{2k+1} meets V_k^n in a V_h^n which is the locus of a single infinity of $(h-1)$-spaces. Now if we let V_h^n be projected upon an S_{2h-i}, $(i=0, 1, \cdots, h-1)$, of S_{k+h+1}, then we have for projection an h-dimensional variety of order n with a double i-dimensional variety of order b_{h-i} and an $(i-1)$-dimensional pinch variety of order j_{h-1} lying on the double variety. If $i=0$, the projection in S_{2h} has b_h improper double points.

Suppose $h=k$, and then we have the given V_k^n itself. A general S_{2k-i}-projection of this V_k^n contains a double i-dimensional variety of order b_{k-i} upon which lies an $(i-1)$-dimensional variety of order j_{k-i}.

In order to determine b_k we find it necessary to determine b_h. This determination will be much facilitated if we make use of the two following results already known.

(A) The characteristics $b_0, b_1, \cdots, b_k; j_0, j_1, \cdots, j_{k-1}$ of any V_k^n in r-spaces satisfy the relations*

$$\begin{aligned} 2b_h &= 2b_{h-1} - j_{h-1} \\ &= 2b_{h-2} - j_{h-1} - j_{h-2} \\ &\quad \cdots\cdots\cdots\cdots \\ &\quad \cdots\cdots\cdots\cdots \end{aligned}$$

* B. C. Wong, *On certain characteristics of k-dimensional varieties in r-space*, this Bulletin, vol. 38, pp. 725–730. The notations used in this paper are slightly different from those adopted in the present work.

$$= 2b_1 - j_{h-1} - j_{h-2} - \cdots - j_1$$
$$= 2b_0 - j_{h-1} - j_{h-2} - \cdots - j_1 - j_0.$$

Here b_0 is to be taken equal to $n(n-1)/2$* and j_0 is the rank of the curve C^n common to V_k^n and an S_{k+2}.†

(B) The number of pinch points on the double curve of a V_k^n which is the locus of a single infinity of S_{h-1}'s in an S_{2h-1} is‡

$$j_{h-1} = 2(n - h + hp).$$

Combining these two results, we find that

$$b_h = b_0 - (1/2)\sum_{i=0}^{h-1} j_i$$
$$= (n - h)(n - h - 1)/2 - k(k + 1)p/2.$$

If $h=1$ and 2, we have formulas (1) and (2), respectively. For $h=k$, we have

$$b_k = (n - k)(n - k - 1)/2 - k(k + 1)p/2$$

as the number of apparent double points on a V_k^n which is the locus of a single infinity of S_{k-1}'s in an S_{2k+1} and this is the number it was our purpose to determine.

THE UNIVERSITY OF CALIFORNIA

---

* If we define b_h as the number of secant lines of a V_k^n of an S_{2h+2} that meet a given line of S_{2h+2}, we see that b_0 is the number of lines determined by n given points in a plane.

† We may define j_{h-1} as the number of tangent lines of a V_k^n of an S_{2h} that pass through a given point of S_{2h}. Then, j_0 is the class of a plane curve which is the plane projection of the curve C^n of intersection of V_k and an S_{k+2}.

‡ B. C. Wong, *On the number of stationary tangent S_{t-1}'s to a V_k in an S_{tk+k-1}*, this Bulletin, vol. 39, pp. 608–610.

given value of k. Then the formula for b_k must be a function of n and p, and of k also.

Consider the ruled surface F^n in which V_k^n is met by a general S_{k+3} of S_{2k+1}. The projection of F^n, if it is in an S_5, has b_2, given by formula (2), apparent double points; and, if it is in an S_3, has a double curve of order b_1 given by formula (1). On this double curve lie a finite number, j_1, of pinch points. This number is known and will be given subsequently.

Next, consider the planed variety V_3^n common to V_k^n and a general S_{k+4} of S_{2k+1}. The projection in an S_7 of this V_3^n has b_3 apparent double points. Projecting this projection successively upon an S_6, an S_5, and an S_4, we see that the resulting variety in S_6 has b_3 improper double points; that the one in S_5 contains a double curve of order b_2 upon which lie j_2 pinch points; and, finally, that the one in S_4 contains a double surface of order b_1 upon which lies a pinch curve of order j_1.

In general, an $S_{k+h+1}(k \geqq h > 0)$ of S_{2k+1} meets V_k^n in a V_h^n which is the locus of a single infinity of $(h-1)$-spaces. Now if we let V_h^n be projected upon an S_{2h-i}, $(i=0, 1, \cdots, h-1)$, of S_{k+h+1}, then we have for projection an h-dimensional variety of order n with a double i-dimensional variety of order b_{h-i} and an $(i-1)$-dimensional pinch variety of order j_{h-1} lying on the double variety. If $i=0$, the projection in S_{2h} has b_h improper double points.

Suppose $h=k$, and then we have the given V_k^n itself. A general S_{2k-i}-projection of this V_k^n contains a double i-dimensional variety of order b_{k-i} upon which lies an $(i-1)$-dimensional variety of order j_{k-i}.

In order to determine b_k we find it necessary to determine b_h. This determination will be much facilitated if we make use of the two following results already known.

(A) The characteristics $b_0, b_1, \cdots, b_k; j_0, j_1, \cdots, j_{k-1}$ of any V_k^n in r-spaces satisfy the relations*

$$\begin{aligned} 2b_h &= 2b_{h-1} - j_{h-1} \\ &= 2b_{h-2} - j_{h-1} - j_{h-2} \\ &\quad \cdots\cdots\cdots\cdots \\ &\quad \cdots\cdots\cdots\cdots \end{aligned}$$

* B. C. Wong, *On certain characteristics of k-dimensional varieties in r-space*, this Bulletin, vol. 38, pp. 725–730. The notations used in this paper are slightly different from those adopted in the present work.

$$= 2b_1 - j_{h-1} - j_{h-2} - \cdots - j_1$$
$$= 2b_0 - j_{h-1} - j_{h-2} - \cdots - j_1 - j_0.$$

Here b_0 is to be taken equal to $n(n-1)/2$* and j_0 is the rank of the curve C^n common to V_k^n and an S_{k+2}.†

(B) The number of pinch points on the double curve of a V_k^n which is the locus of a single infinity of S_{h-1}'s in an S_{2h-1} is‡

$$j_{h-1} = 2(n - h + hp).$$

Combining these two results, we find that

$$b_h = b_0 - (1/2)\sum_{i=0}^{h-1} j_i$$
$$= (n - h)(n - h - 1)/2 - k(k + 1)p/2.$$

If $h=1$ and 2, we have formulas (1) and (2), respectively. For $h=k$, we have

$$b_k = (n - k)(n - k - 1)/2 - k(k + 1)p/2$$

as the number of apparent double points on a V_k^n which is the locus of a single infinity of S_{k-1}'s in an S_{2k+1} and this is the number it was our purpose to determine.

THE UNIVERSITY OF CALIFORNIA

---

* If we define b_h as the number of secant lines of a V_k^n of an S_{2h+2} that meet a given line of S_{2h+2}, we see that b_0 is the number of lines determined by n given points in a plane.

† We may define j_{h-1} as the number of tangent lines of a V_k^n of an S_{2h} that pass through a given point of S_{2h}. Then, j_0 is the class of a plane curve which is the plane projection of the curve C^n of intersection of V_k and an S_{k+2}.

‡ B. C. Wong, *On the number of stationary tangent S_{t-1}'s to a V_k in an S_{tk+k-1}*, this Bulletin, vol. 39, pp. 608–610.

QUADRATIC RESIDUES IN FACTORIZATION*

BY MARSHALL HALL

1. *Introduction.* The purpose of this paper is to establish a certain theorem which is useful in the factorization of large numbers. Quadratic residues have been frequently used in factorization, particularly by M. Kraïtchik in volume II of his *Recherches sur la Théorie des Nombres.* Beeger† has proved several propositions on the use of quadratic residues in factorization which Kraïtchik tacitly assumed. Quadratic forms are the most convenient representations of a number which give material information as to the type of its prime factors. The knowledge of several quadratic residues of a number is of great aid in finding its factors, but in identifying a prime by its quadratic residues our proof is negative, in that the same result might come about through an error in the calculation. It is to eliminate much of the calculation involved in identifying a prime, and to make the proof of primality positive in character, that the present paper has been undertaken.

2. *Definition of Apparent Residues and Non-Residues.* Following Kraïtchik,‡ I define (quadratic) apparent residues and apparent non-residues in the following manner. If a, b are odd primes >1, if $(N/\alpha)=+1$ and if $a'=(-1/a)a$, then a' is said to be an apparent residue of N.

If $(N/b)=-1$ and if $b'=(-1/b)b$, then b' is said to be an apparent non-residue of N. According as the Jacobian symbol $(-1/N)$ is $+1$ or -1, -1 is said to be an apparent residue or non-residue of N. Similarly the apparent characters of $+2$ and -2 with respect to N are defined.

We define compound apparent residues and non-residues by calling the product of two apparent residues or two apparent non-residues an apparent residue, and the product of an apparent residue and an apparent non-residue an apparent non-

* Presented to the Society, March 25, 1932.

† Nieuw Archief voor Wiskunde, (2), vol. 16, No. 4, pp. 37–42.

‡ *Recherches sur la Théorie des Nombres*, vol. 2, 1929, p. 8. For Kraïtchik's "résidu éventuel" I write "apparent residue."

residue. No apparent character is given to a number not relatively prime to N, as it is our purpose to find the factors of N, which is supposed to be of unknown composition.

If a number N be prime, then its apparent residues are true residues (that is, $x^2 \equiv a \pmod{N}$, a any apparent residue) and the product of any two apparent non-residues, is a true residue.

3. *A Preliminary Theorem.* We shall now establish various preliminary theorems on apparent residues. We show first the following result.

THEOREM 1. *A number which is a quadratic residue of every prime not dividing it is a perfect square.*

As every odd power of 2 is a non-residue of 3, we may assume that such a number contains an odd factor. Let N be a quadratic residue of every prime not dividing it, and $N = j^2 n$, where n contains no square factors. Then n is a quadratic residue of the same primes as N.

Let $n = p_1 p_2 \cdots p_r$ and let furthermore a be a quadratic non-residue of p_1 (p_1 odd) and b_i be a quadratic residue of p_i, $(i = 2, \cdots, r)$. The congruences $x \equiv 1 \pmod 4$, $x \equiv a \pmod{p_1}$,

$$x \equiv b_i \pmod{p_i} \equiv 1 \pmod 8 \text{ if } p_i = 2, \quad (i = 2, \cdots, r),$$

then always have a solution s, as the moduli are relatively prime. There are an infinite number of prime values for the general solution $x = 4kn + s$. Choose one not dividing j. Then

$$\left(\frac{x}{p_1}\right) = \left(\frac{p_1}{x}\right) = -1, \left(\frac{x}{p_i}\right) = \left(\frac{p_i}{x}\right) = +1, \quad (i = 2, \cdots, r);$$

hence $(N/x) = (n/x) = -1$, contrary to hypothesis, and N must be a perfect square. We have actually proved more than the theorem, namely, that if N is a quadratic residue of all primes except a finite number, then N must be a perfect square.

4. *Compound Characters.* We shall now show that the product of two prime apparent non-residues of a number N is equivalent to a prime apparent residue.

THEOREM 2. *There exists a prime apparent residue of N which has the same quadratic character with respect to any factor of N as the product of the two given prime apparent non-residues.*

Given p_1 and p_2, two prime apparent non-residues of N; each is then a non-residue of an odd number of prime factors of N. There exists a prime $p=kN+p_1p_2$ which is such that $p\equiv 1 \pmod 4$. This is a non-residue of an even number of prime factors of N, and is consequently an apparent residue of N. For if N has s prime factors, $N=f_1f_2\cdots f_s$ (f's not necessarily distinct), we may arrange the f's in the following way.

Class I. $\left(\frac{p_1}{f_i}\right)=\left(\frac{p_2}{f_i}\right)=+1, \qquad (i=1, 2, \cdots, \lambda).$

Class II. $\left(\frac{p_1}{f_i}\right)=-1, \left(\frac{p_2}{f_i}\right)=+1, \qquad (i=\lambda+1, \lambda+2, \cdots, \mu).$

Class III. $\left(\frac{p_1}{f_i}\right)=+1, \left(\frac{p_2}{f_i}\right)=-1, \qquad (i=\mu+1, \mu+2, \cdots, \nu).$

Class IV. $\left(\frac{p_1}{f_i}\right)=\left(\frac{p_2}{f_i}\right)=-1, \qquad (i=\nu+1, \nu+2, \cdots, s).$

Here p is a residue of Classes I and IV (that is, a residue of every f in these classes), and a non-residue of Classes II and III, that is, a non-residue of $\nu-\lambda$ factors of N. The prime p_1 is a non-residue of Classes II and IV, that is, $\mu-\lambda+s-\nu$ factors; p_2 is a non-residue of Classes III and IV, that is, $s-\mu$ factors, and $\nu-\lambda\equiv s-\mu+\mu-\lambda+s-\nu \pmod 2$. As $s-\mu$ and $\mu-\lambda+s-\nu$ are odd, then $\nu-\lambda$ is even and p is an apparent residue of N. Also $p\equiv p_1p_2 \pmod{f_i}$ for any prime factor of N and has the same quadratic character as the product for any factor.

In like manner, it may be shown that the product of two prime apparent residues is equivalent to a prime apparent residue, and the product of a prime apparent residue and a prime apparent non-residue is equivalent to a prime apparent non-residue. Consequently, we may treat compound apparent residues and non-residues as if they were the equivalent primes. In this connection, it may be worth while to mention that a prime $p\equiv 3 \pmod 4$ is considered compound. For example, $+3=-1\cdot -3$ is a compound apparent residue of 35.

5. *Uniqueness of Apparent Residues.*

THEOREM 3. *If the apparent residues of a number A are included in the apparent residues of a number B, not a square, both their apparent residues and their apparent non-residues will coincide.*

For if b is an apparent residue of B, but not of A, and c is an apparent non-residue of both, then bc is an apparent residue of A, but not of B. Such a c must exist, for if all the non-residues of A were residues of B, then B would be a quadratic residue of every prime except a finite set (divisors of A or B) and B would be a perfect square, contrary to the assumption. Hence the product AB is a square, and consequently, aside from squared factors, a number is uniquely determined by its prime apparent residues.

6. *A Test for Primality.* We are now in a position to prove Theorem 4.

THEOREM 4. *If all the apparent residues of a number N are true residues, then the number N is either a prime or a power of a prime.*

PROOF. Let $R=r_1, r_2, r_3, \cdots$ be the set of prime apparent residues of the given number N, and $S=s_1, s_2, \cdots$ be the set of prime apparent non-residues. The sets R and S contain all primes except those dividing N. All primes of the set R are by definition true residues of N. As by Theorem 2 any product $s_i s_j$ has the same character with respect to every factor of N as some r, that is, a residue, if we hold s_i fixed and let s_j vary, we see that the quadratic character, with respect to any prime factor of N, of all s's is determined by a single one.

Then for any prime factor p of N we have the following result. Every prime in R is a residue of p and either (1) all primes of S are non-residues of p, or (2) all primes of S are residues of p.

In the first case, a single *prime* is determined by Theorem 3. In the second case, there is no prime at all, as this is the characterization of a square by Theorem 1. Hence N is divisible only by a single prime, and the theorem follows.

7. *Applications.* In an actual problem of factorization, it is impracticable to determine the true quadratic character of every number. But if all the apparent residues considered are shown to be true residues, the problem is essentially this.

In the two sets

$$A = a_1, a_2, a_3, \cdots, a_r, \qquad B = b_1, b_2, b_3, \cdots, b_s,$$

A and B contain all primes less than some p (-1 considered a prime), we know that every a_i is a true residue, and that every

b_ib_j is a true residue of the number N under consideration. We wish to know what restriction this information places upon possible factors of N.

8. *A Practicable Test for Primality.* We define L_p as the least number, not a square, which is a quadratic residue of all primes $\leqq p$. In other terms, every number less than L_p which has

$$-1, 2, -3, \cdots, (-1/p)p$$

as apparent residues must be a perfect square. (Numbers with factors less than p are excluded by the definition of apparent residues.)

THEOREM 5. *If all the factors (not necessarily prime factors) of N lie below L_p, and if $-1, 2, \cdots, (-1/p)p$ may be divided into two classes $A(a_1, a_2, \cdots, a_r)$, the apparent residues of N, and $B(b_1, b_2, \cdots, b_s)$, the apparent non-residues of N, such that every a_i is a true residue of N, and every b_ib_j is a true residue, then N is either a prime or a power of a prime.*

PROOF. For any prime factor q of N we have one of two cases:

CASE I. Every prime in A is a residue of q and every prime in B is a non-residue of q.

CASE II. Every prime in A is a residue of q and every prime in B is a residue of q.

Case II is to be rejected as $-1, 2, \cdots, (-1/p)p$ would be apparent residues of q, and q lies below L_p. This would make q a square, though also a prime.

If q_1 is a prime divisor of N belonging to Case I, then $(q_1/p_j) = (N/p_j)$ for p_j any of $-1, 2, \cdots, (-1/p)p$. In this event $([N/q_1]/p_j) = +1$ and as $N/q_1 < L_p$, it follows that N/q_1 must be a square. Hence q_1 is contained in N to an odd power. If q_2 were another prime dividing N, we would also have N/q_2 a square though divisible by an odd power of q_1. Hence N must be either a prime or the power of a prime.

The numbers L_p seem to increase rapidly as p increases. Lehmer* has found the values up to L_{61} which is 48,473,881. By Theorem 1 it follows that L_p tends to infinity with p.

An asymptotic formula for L_p or some lower bound, such as $2^{p/2}$, would be quite valuable in applying this last theorem.

* American Mathematical Monthly, vol. 35, p. 121. See also Kraitchik, *Recherches sur la Théorie des Nombres*, vol. 1, p. 46.

EXAMPLE. $N=22,253,377$. We find the following quadratic forms.

1. $N = (2361)^2 + (4084)^2.$
2. $N = (3467)^2 + 2(2262)^2.$
3. $N = (3905)^2 + 3(1528)^2.$
4. $N = (4703)^2 + 2^{12}\cdot 3\cdot 11.$
5. $N = (4714)^2 + 3^2\cdot 11^2\cdot 29.$
6. $N = (4736)^2 - 3^2\cdot 11\cdot 13\cdot 137.$
   $N = (4717)^2 + 2^3\cdot 3\cdot 137.$
7. $N = (4723)^2 - 2^3\cdot 3^3\cdot 13\cdot 19.$
8. $N = (4777)^2 - 2^4\cdot 3^4\cdot 19\cdot 23.$

This gives us the set $A(-1, 2, -3, -11, 13, -19, -23, 29)$ and shows them to be true residues. For the set B we find:

9. $5N = (10,525)^2 + 2^2\cdot 11^2\cdot 5\cdot 7\cdot 29.$
10. $3N = (8174)^2 - 7^2\cdot 5\cdot 13\cdot 17.$
11. $3N = (7919)^2 + 2\cdot 5\cdot 7\cdot 17\cdot 41\cdot 83.$
    $N = (4722)^2 - 23^2\cdot 83.$
12. $3N = (8206)^2 - 5\cdot 7\cdot 13\cdot 31\cdot 41.$
13. $5N = (10,533)^2 + 2^2\cdot 17\cdot 47\cdot 101.$
    $N = (4705)^2 + 2^7\cdot 3^2\cdot 101.$
14. $5N = (10,547)^2 + 2^2\cdot 11\cdot 17\cdot 37.$

This gives us the set $B(5, -7, 17, -31, 37, 41, -47)$, and b_ib_j is a residue. As N is not divisible by 2 or 3, any factor must lie below 9,257,329 $= L_{47}$; hence, if of Case II type, must be a perfect square. Consequently N is a prime or a power of a prime. As $N\equiv 2 \pmod 5$ and $\equiv 3 \pmod{13}$, it is neither a square nor a cube. Having no factor under 50, it cannot be a higher power. Hence N is identified as a prime.

TRINITY COLLEGE, CAMBRIDGE

A COMPLETE CENSUS OF 4×4 MAGIC SQUARES*

BY D. N. LEHMER

A complete enumeration of magic squares of order 3 is easy to make and the result turns out to be 72. In this enumeration two squares are considered different if they do not have the same numbers in corresponding cells. In this and in the study of 4×4 squares we define a magic square as one in which the elements in every row and column add up to the same sum, this sum being given for the square of order n by the formula

$$\frac{n(n^2+1)}{2}.$$

With this definition a magic square remains magic after any permutation of the rows among themselves, or of the columns among themselves, or by an interchange of rows and columns. This associates with any magic square a set of $2(n!)^2$ squares all of which are magic.

By a permutation of the rows and columns of any square the largest entry in the square may be transferred to any particular cell; in particular it may be placed in the lower left hand corner. The other rows and columns may then be permuted so that the entries in the bottom line and in the left hand column read in descending order of magnitude. Also by an interchange of rows and columns, if necessary, the element in the bottom row next to the left hand corner may be made larger than the element in the left hand column which is next above the left hand corner. No further adjustments are then possible and the square will be said to be *normalized*. Thus the square

22	5	8	11	19		3	21	19	12	10
4	7	15	18	21		6	4	22	20	13
6	14	17	25	3	becomes	14	7	5	23	16
13	16	24	2	10		17	15	8	1	24
20	23	1	9	12		25	18	11	9	2

after normalization;

* Presented to the Society, November 29, 1929, and June 23, 1933.

It is clear that two normalized squares which do not agree entry for entry can not be transformed one into the other by any permutations of rows or of columns, or by interchange of rows and columns. To find the total number of squares of order n all we need to find is the number of normalized squares. This number, multiplied by $2(n!)^2$ will give the total number of squares for that order.

To exhibit the method used in the discussion of 4×4 squares we derive the normalized squares of order three. The sum in each row and column must be 15. We assume the entry 9 in the lower left hand corner. The other two entries in the bottom row and in the left hand column must add to $15-9=6$. The only partitions of 6 obtainable from the numbers 1, 2, 3, 4, 5, 6, 7, 8, the numbers in the partition being distinct, are $5+1$ and $4+2$. For the normalized square, therefore, we must write the bottom row 9, 5, 1 and the left hand column 9, 4, 2. This leaves only the numbers 3, 6, 7, 8 with which to complete the square. The two remaining numbers in the second column must add to 10, and the two in the second row to 11. The only partition of 10 available is $7+3$ and the only partition 11 available is $8+3$. This indicates that the cell in which these two lines intersect must be filled with the number 3. We get thus the single normalized square of order 3 to be

$$\begin{array}{ccc} 2 & 7 & 6 \\ 4 & 3 & 8 \\ 9 & 5 & 1 \end{array}$$

All other magic squares of order 3 are therefore obtainable from this single one by permutation of its rows and columns and by interchanging its rows and columns. The number of such transformations being $2(3!)^2=72$, the total number of magic squares of order 3 is therefore 72.

For the 4×4 squares the magic sum is 34. After placing 16 in the lower left hand corner the sum of the three remaining numbers in the first row and in the first column must be $34-16=18$. There are 19 partitions of 18 using the numbers from 1 to 15, if the numbers in the partitions are assumed distinct as they are in this case. They are (15, 2, 1), (14, 3, 1), (13, 4, 1), (12, 5, 1), (11, 6, 1), (10, 7, 1), (9, 8, 1), (13, 3, 2), (12, 4, 2), (11, 5, 2), (10, 6, 2), (9, 7, 2), (11, 4, 3), (10, 5, 3), (9, 6, 3), (8, 7, 3),

(9, 5, 4), (8, 6, 4) and (7, 6, 5). Every normalized magic square must then have one of these *triads* for the numbers (besides the number 16) in its first row and in its first column. Moreover having selected a triad for the bottom row the triad for the first column must be made up of numbers distinct from those in the triad first selected, and the largest number in it must be smaller than the largest number in the chosen triad. Thus starting with the triad (15, 2, 1) all the above triads which contain 1 or 2 are not available and there are only seven left, namely, (11, 4, 3), (10, 6, 2), (9, 6, 3), (8, 7, 3), (8, 6, 4), (7, 6, 5). Similarly for the other choices of the triad in the bottom row. It is observed that no triad for the column can be used with a triad chosen for the bottom row if it precedes it in the above list. Thus, for example, for the triad (10, 7, 1), as a choice for the bottom row we can use only three triads (9, 6, 3), (9, 5, 4), (8, 6, 4), and for the triad (9, 5, 4) there is only one available: (8, 7, 3).

These combinations of triads in the bottom line with triads in the first column (67 in all) must each be examined for possible magic squares. The method is analogous to the one used in the above discussion for 3×3 squares and will be clear from the example of the first case in which we have the triad (15, 2, 1) for the bottom row and (11, 4, 3) for the first column. These triads use up the numbers 1, 2, 3, 4, 11, 15 and 16. The remaining numbers are 5, 6, 7, 8, 9, 10, 12, 13, 14. Out of these a triad must be selected for the second column which shall give a sum $34-15=19$. The only such is (8, 6, 5). A triad must also be selected for the second row with a sum 23. The only ones are (12, 6, 5), (10, 8, 5), (10, 7, 6), (9, 8, 6). But the vertical column and the horizontal row must have one and only one element in common. This rules out of consideration the 1st, 2d and 4th of the above triads, and we have left only the triads (8, 6, 5) and (10, 7, 6) to consider, with the common element 6 which must therefore be located in the intersection cell. The other elements must then be suitably placed and the result is the two squares:

3	8	9	14			3	5	12	14
4	5	13	12		and	4	8	13	9
11	6	10	7			11	6	7	10
16	15	2	1			16	15	2	1

The discussion for the other cases may be made in the same way. The following table gives for each bottom line triad the set of triads available for the first column, the subscript indicating the number of squares obtained from that particular choice of triads.

Bottom line triads	Available first column triads
(15, 2, 1)	$(11, 4, 3)_2$, $(10, 5, 3)_{12}$, $(9, 6, 3)_{11}$, $(9, 5, 4)_{11}$, $(8, 7, 3)_2$, $(8, 6, 4)_3$, $(7, 6, 5)_4$.
(14, 3, 1)	$(12, 4, 2)_3$, $(11, 5, 2)_{13}$, $(10, 6, 2)_7$, $(9, 7, 2)_{13}$, $(9, 5, 4)_{20}$, $(8, 6, 4)_2$, $(7, 6, 5)_1$.
(13, 4, 1)	$(11, 5, 2)_{14}$, $(10, 6, 2)_2$, $(10, 5, 3)_{10}$, $(9, 7, 2)_{11}$, $(9, 6, 3)_{14}$, $(8, 7, 3)_5$, $(7, 6, 5)_2$.
(13, 3, 2)	$(12, 5, 1)_{14}$, $(11, 6, 1)_{15}$, $(10, 7, 1)_{11}$, $(9, 8, 1)_{22}$, $(9, 5, 4)_7$, $(8, 6, 4)_5$, $(7, 6, 5)_0$.
(12, 5, 1)	$(11, 4, 3)_8$, $(10, 6, 2)_3$, $(9, 7, 2)_{20}$, $(9, 6, 3)_{13}$, $(8, 7, 3)_3$, $(8, 6, 4)_2$.
(12, 4, 2)	$(11, 6, 1)_4$, $(10, 7, 1)_6$, $(10, 5, 3)_4$, $(9, 8, 1)_2$, $(9, 6, 3)_1$, $(8, 7, 3)_0$, $(7, 6, 5)_0$.
(11, 6, 1)	$(10, 5, 3)_{14}$, $(9, 7, 2)_{15}$, $(9, 5, 4)_{14}$, $(8, 7, 3)_5$.
(11, 5, 2)	$(10, 7, 1)_{14}$, $(9, 8, 1)_{14}$, $(9, 6, 3)_5$, $(8, 7, 3)_0$, $(8, 6, 4)_5$.
(11, 4, 3)	$(10, 7, 1)_3$, $(10, 6, 2)_1$, $(9, 8, 1)_4$, $(9, 7, 2)_0$, $(7, 6, 5)_0$.
(10, 7, 1)	$(9, 6, 3)_{14}$, $(9, 5, 4)_{11}$, $(8, 6, 4)_1$.
(10, 6, 2)	$(9, 8, 1)_2$, $(9, 5, 4)_2$, $(8, 7, 3)_1$.
(10, 5, 3)	$(9, 8, 1)_{13}$, $(9, 7, 2)_7$, $(8, 6, 4)_0$.
(9, 8, 1)	$(7, 6, 5)_8$.
(9, 7, 2)	$(8, 6, 4)_7$.
(9, 5, 4)	$(8, 7, 3)_6$.

It is seen that there are in all 468 normalized squares of order 4, and multiplying this number by $2(4!)^2 = 1152$, we get for the total number of magic squares of order 4 the number 539,136.

THE UNIVERSITY OF CALIFORNIA

ON THE THEORY OF FOURIER TRANSFORMS

BY EINAR HILLE AND J. D. TAMARKIN

1. *Introduction.* Let $g(s) \subset L_2$ over $(-\infty, \infty)$. Let

$$G(u; a) = (2\pi)^{-1/2} \int_{-a}^{a} e^{-ius} g(s) ds.$$

According to the classical result of the Plancherel theory of Fourier transforms, $G(u; a)$ tends in the mean of order 2 to a function $G(u) \subset L_2$ as $a \to \infty$. This function is designated as the Fourier transform (in L_2) of $g(s)$. We shall write

$$T\{u; g(s)\} = G(u) = \operatorname*{l.i.m}_{a \to \infty} G(u; a). \tag{1}$$

The functions $g(s)$ and $G(u)$ are reciprocal in the sense that

$$g(s) = T\{s; G(-u)\}, \tag{2}$$

which means that

$$g(s) = \operatorname*{l.i.m}_{a \to \infty} g(s; a); \quad g(s; a) = (2\pi)^{-1/2} \int_{-a}^{a} e^{ius} G(u) du. \tag{3}$$

As an immediate consequence of the convergence in the mean of $G(u; a)$ and $g(s; a)$, we have, almost everywhere,

$$G(u) = (2\pi)^{-1/2} \frac{d}{du} \int_{-\infty}^{\infty} g(s) \frac{1 - e^{-ist}}{is} ds, \tag{4}$$

$$g(s) = (2\pi)^{-1/2} \frac{d}{ds} \int_{-\infty}^{\infty} G(u) \frac{1 - e^{ius}}{-iu} du. \tag{5}$$

The reciprocity between g and G is expressed here in terms not involving the convergence in the mean.

Assume now that $g(s) \subset L_p$, $1 < p < 2$. Denote by p' the *conjugate* exponent, $p' = p/(p-1)$, $1/p + 1/p' = 1$. Titchmarsh* showed that Plancherel's theory can be extended, at least in part, to the present case. Indeed he proved that $G(u; a)$ con-

* E. C. Titchmarsh, *A contribution to the theory of Fourier transforms*, Proceedings of the London Mathematical Society, (2), vol. 23 (1925), pp. 279–289. We have slightly modified Titchmarsh's notation inasmuch as he deals with cosine- and sine-transforms, while we use exponential transforms.

verges in the mean of order p' to a function $G(u) \subset L_{p'}$, which satisfies the same reciprocity relations (4), (5). Titchmarsh did not prove, however, that, conversely, (3) holds. On the basis of the material developed by Titchmarsh we do not know whether $g(s; a)$ tends in the mean of order p to $g(s)$; hence we do not know whether $g(s)$ is the Fourier transform (in L_p) of $G(-u)$. A scrutiny of the literature reveals the unexpected fact that this question has never been investigated,* which, in our opinion, represents an undesirable gap in the theory of Fourier transforms. The purpose of the present note is to fill in this gap and to prove (3) in the case $1 < p \leqq 2$. An analogy between the theory of Fourier transforms and that of Fourier series should be pointed out here. We may consider the Fourier transform $G(u)$ of $g(s)$ as an analog of the *sequence* of Fourier coefficients $\{G_n\}$ of $g(s)$ (in the case where g is periodic and is being expanded in Fourier series). Then $g(s; a)$ appears as an analog of the nth partial sum $s_n(s; g)$ of $g(s)$ which is known to converge to $g(s)$ in the mean of order p. This analogy goes even further. It is known that if $g(s) \subset L_1$, over $(-\pi, \pi)$, $s_n(s; g)$ does not necessarily converge in the mean of order 1 to $g(s)$.† We shall prove that this is also the case for the Fourier transforms.

It will be essential for our discussion to consider the class of functions $\subset L_p$ over $(-\infty, \infty)$ as a linear vector metric complete space with the metric defined by

$$\|g\|_p = \left[\int_{-\infty}^{\infty} |g(s)|^p ds\right]^{1/p}.$$

By $\mathfrak{L}^p$ we shall mean the limit in the mean of order p. Some known properties of linear transformations of such spaces will be used in the sequel.‡

* Contrary to our unfounded statement in a recent paper. See E. Hille and J. D. Tamarkin, *On the summability of Fourier series*, III, Mathematische Annalen, vol. 108 (1933), pp. 525–577 (p. 530). According to results of the present note, however, this statement is correct.

† The case $p=1$ goes back to H. Hahn, *Ueber die Darstellung gegebener Funktionen durch singuläre Intergrale*, II, Denkschriften Wiener Akademie, Math.-Nat. Kl., vol. 93 (1916), p. 557–692 (p. 681). For $p>1$ see M. Riesz, *Sur les fonctions conjuguées*, Mathematische Zeitschrift, vol. 27 (1927), pp. 218–244 (p. 230). M. Riesz states at the end of his paper that to his results concerning Fourier series there correspond similar results for Fourier integrals.

‡ We refer for these properties to the recent book by S. Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932.

2. *The Case* $1<p\leqq 2$. If $g(s)\subset L_p$, $1<p\leqq 2$, we have

$$G(u) = \operatorname*{l.i.m.}_{a\to\infty} (2\pi)^{-1/2}\int_{-a}^{a} g(t)e^{-iut}dt.$$

Here we may multiply by any function $\subset L_p$ and integrate under the integral sign. Thus

$$\begin{aligned} g(s;a) &= (2\pi)^{-1/2}\int_{-a}^{a} G(u)e^{isu}du \\ &= (2\pi)^{-1}\int_{-\infty}^{\infty} g(t)dt\int_{-a}^{a} e^{iu(s-t)}du \qquad (6) \\ &= \frac{1}{\pi}\int_{-\infty}^{\infty} g(t)\frac{\sin a(s-t)}{s-t}dt. \end{aligned}$$

Let $f(s)$ be any function $\subset L_p$. The integral

$$\tilde{f}(s) = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{f(t)}{s-t}dt$$

exists (in the sense of Cauchy's principal value at $t=s$) almost everywhere and is designated as the function conjugate to $f(s)$.

Moreover, $\tilde{f}(s)\subset L_p$ whenever $f(s)\subset L_p$. Furthermore, there exists a positive constant M_p depending only on p and such that*

$$\|\tilde{f}\|_p \leqq M_p\|f\|_p. \qquad (7)$$

Upon introducing the functions

$$g_a'(t) = g(t)\cos at, \qquad g_a''(t) = -g(t)\sin at$$

and their conjugates $\tilde{g}_a'(s)$, $\tilde{g}_a''(s)$,† we may rewrite (6) in the form

$$g(s;a) = \tilde{g}_a'(s)\sin as + \tilde{g}_a''(s)\cos as.$$

By (7),

$$\|\tilde{g}_a'\|_p \leqq M_p\|g(t)\cos at\|_p \leqq M_p\|g\|_p,\ \|\tilde{g}_a''\|_p \leqq M_p\|g\|_p.$$

---

* M. Riesz, loc. cit., p. 234.

† For the analogous procedure in the case of Fourier series, see M. Riesz, loc. cit., p. 230, and A. Kolmogoroff, *Sur les fonctions harmoniques conjuguées et sur les séries de Fourier*, Fundamenta Mathematicae, vol. 7 (1925), pp. 23–28.

Hence

$$\|g(s;a)\|_p \leqq \|\tilde{g}_a'\|_p + \|\tilde{g}_a''\|_p \leqq 2M_p\|g\|_p. \tag{8}$$

For each a, the expression $g(s;a)$ constitutes as a linear transformation on L_p to L_p, and (8) shows that the family of these transformations, obtained when a varies, is uniformly limited.

Now assume that $g(s)$ is a step function of a finite number of steps, vanishing outside of a finite interval $(-N, N)$. For such a function we shall prove directly that

$$\|g(s;a) - g(s)\|_p \to 0 \text{ as } a \to \infty.$$

Since our function $g(s)$ is a linear combination of step functions defined by

$$g(s) = \begin{cases} 1 \text{ if } \alpha \leqq s \leqq \beta, \\ 0 \text{ outside } (\alpha, \beta), \end{cases} \qquad (-N \leqq \alpha < \beta \leqq N),$$

it will be sufficient to prove our assertion for a function of this type. For such a function, however,

$$g(s;a) = \int_\alpha^\beta \frac{\sin a(s-t)}{s-t} dt.$$

It is readily seen that

$$g(s;a) = O[(s-N)^{-1}a^{-1}], \qquad (s \geqq 2N).$$

Since $g(s) = 0$ for $|s| > N$, we have

$$\int_{2N}^\infty |g(s) - g(s;a)|^p ds = O\left[a^{-p}\int_{2N}^\infty \frac{ds}{(s-N)^p}\right] = O(a^{-p})$$

as $a \to \infty$. Similarly

$$\int_{-\infty}^{-2N} |g(s) - g(s;a)|^p ds = O(a^{-p}).$$

On the other hand the integral

$$\frac{1}{\pi}\int_\alpha^\beta \frac{\sin a(s-t)}{s-t} dt = \frac{1}{\pi}\int_{-2N}^{2N} g^*(t)\frac{\sin a(s-t)}{s-t} dt$$

may be considered as the classical Dirichlet integral of the function $g^*(s)$ which is periodic, of period $4N$, and coincides with $g(s)$ in the interval $(-2N, 2N)$. It is well known from the theory

of Fourier series, and also can be proved directly, that this Dirichlet integral converges to $g^*(s)$ in the mean of order p, over the interval $(-2N, 2N)$. Hence

$$\int_{-2N}^{2N} |g(s) - g(s; a)|^p ds$$

$$= \int_{-2N}^{2N} |g^*(s) - g(s; a)|^p ds \to 0 \text{ as } a \to \infty .$$

Consequently

$$\int_{-\infty}^{\infty} |g(s) - g(s; a)|^p ds = \int_{-2N}^{2N} \cdots + \int_{-\infty}^{-2N} \cdots + \int_{2N}^{\infty} \cdots \to 0$$

as $a \to \infty$, which is the desired result.

Since the family of transformations represented by $g(s; a)$ is uniformly limited, and since the step-functions constitute a dense sub-set of L_p, it follows immediately that

$$\|g(s) - g(s; a)\|_p \to 0 \text{ as } a \to \infty ,$$

for an arbitrary function $g(s) \subset L_p$. We may therefore state the following addition to the results of Titchmarsh mentioned above.

THEOREM. *If* $g(s) \subset L_p$, $1 < p \leqq 2$, *and if*

$$G(u) = \operatorname*{l.i.m.}_{a \to \infty}{}^{p'} (2\pi)^{-1/2} \int_{-a}^{a} g(s) e^{-ius} ds$$

is the Fourier transform (in $L_{p'}$*) of* $g(s)$, *then, conversely,* $g(s)$ *is the Fourier transform (in* L_p*) of* $G(-u)$, *and we have*

$$g(s) = \operatorname*{l.i.m.}_{a \to \infty}{}^{p} (2\pi)^{-1/2} \int_{-a}^{a} G(u) e^{ius} du.$$

3. *The Case* $p = 1$. We now proceed to show that our theorem fails in the case $p = 1$. To do this we have to exhibit an example of a function $g(s) \subset L_1$, such that, if we set

$$G(u) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} g(s) e^{-ius} ds,$$

$$g(s; a) = (2\pi)^{-1/2} \int_{-a}^{a} G(u) e^{ius} du,$$

the function $g(s; a)$ does not converge to $g(s)$ in the mean of or-

der 1. The example which we are going to construct will show even more, namely, that the corresponding $g(s;a)$ does not even belong to L_1.

In constructing the example in question we are again guided by the analogy with Fourier series. It is well known* that the series

$$f(x) = \sum_{n=2}^{\infty} \frac{\cos nx}{\log n}$$

is a Fourier series of $f(x) \subset L_1$, but that its partial sums $s_p(x;f)$ do not converge to $f(x)$ in the mean of order 1 over $(-\pi, \pi)$. We now put

$$g(s) = (2\pi)^{-1/2} \int_0^{\infty} \frac{\cos us}{\log (2+u)} du. \tag{9}$$

It is plain that this integral converges uniformly over every interval, finite or not, which is separated from the point $s=0$. Moreover, it is easy to show that $g(s)>0$ for $s\neq 0$. It suffices to consider positive values of s only. Let $s>0$ be fixed. We put

$$q(u) = (2+u)\log^2(2+u), \quad \delta_\nu = \nu\pi/s, \quad (\nu = 0, 1, 2, \cdots).$$

Then, on integrating by parts in the right-hand member of (9), we see that

$$(2\pi)^{1/2} g(s) = \frac{1}{s}\int_0^{\infty} \frac{\sin us}{q(u)} du = \sum_{\nu=0}^{\infty} \frac{1}{s}\int_{\delta_\nu}^{\delta_{\nu+1}} \frac{\sin us}{q(u)} du. \tag{10}$$

Put

$$I_\nu(s) = \frac{1}{s}\int_{\delta_\nu}^{\delta_{\nu+1}} \frac{\sin us}{q(u)} du = (-1)^\nu s^{-2} \int_0^{\pi} \frac{\sin \tau d\tau}{q(\nu\pi + \tau/s)}.$$

Hence

$$|I_\nu| < |I_{\nu-1}|, \qquad (\nu = 1, 2, \cdots),$$

and $\sum_0^{\infty} I_\nu$ appears as an alternating series in which the absolute value of the general term approaches zero and the first term is positive. The sum of such a series is positive. Hence $g(s)>0$. An integration by parts applied to the middle term in (10) shows that, for large values of s, $g(s)=O(s^{-2})$. This ensures the integrability of $g(s)$ over every interval (ϵ, ∞), $\epsilon>0$. Since $g(s)>0$, to

* A. Kolmogoroff, *Sur l'ordre de grandeur de coefficients de la série de Fourier*, Bullétin International de l'Académie Polonaise, Classe de Sciences mathématiques, 1923, pp. 83–86. See also E. Hille and J. D. Tamarkin, *On the summability of Fourier series*, II, Annals of Mathematics, (2), vol. 34 (1933), pp. 329–348 (pp. 347–348).

establish the integrability of $g(s)$ over $(0, \epsilon)$ it suffices to prove that, for any fixed positive α,

$$\int_\epsilon^\alpha g(s)ds$$

is bounded as $\epsilon \to 0$. In view of the uniform convergence of the integral in (9) over $(\epsilon, 1)$ we have

$$\int_\epsilon^1 g(s)ds = (2\pi)^{-1/2} \int_0^\infty \frac{du}{\log(2+u)} \int_\epsilon^1 \cos us\, ds$$

$$= (2\pi)^{-1/2} \left\{ \int_0^\infty \frac{\sin u\, du}{u \log(2+u)} - \int_0^\infty \frac{\sin u\, du}{u \log(2+u/\epsilon)} \right\}.$$

A simple application of the second law of the mean shows that the last integral tends to zero with ϵ. Thus we see that $g(s)$ is integrable over $(0, \infty)$, and, being even, it is integrable over $(-\infty, \infty)$. By the uniqueness theorem of Fourier integrals, it is plain that

$$G(u) = [\log(2 + |u|)]^{-1} = (2\pi)^{-1/2} \int_0^\infty g(s) \cos su\, ds$$

is the Fourier transform of $g(s)$ in L_∞, that is, $G(u; a)$ converges uniformly to $G(u)$ as $a \to \infty$. Now compute

$$g(s; a) = (2\pi)^{-1/2} \int_0^a G(u) \cos su\, du = (2\pi)^{-1/2} \int_0^a \frac{\cos su\, du}{\log(2+u)}.$$

An integration by parts shows that

$$(2\pi)^{1/2} g(s; a) = \frac{\sin sa}{s \log(2+a)} + \frac{1}{s} \int_0^a \frac{\sin su}{q(u)} du.$$

As before, we see that the second term of the right-hand member is $O(s^{-2})$ for $s \to \infty$, and since the first term is not absolutely integrable over $(0, \infty)$ it results that $g(s; a)$ does not belong to L_1.

BROWN UNIVERSITY AND YALE UNIVERSITY

NOTE ON THE LOCATION OF THE CRITICAL POINTS OF GREEN'S FUNCTION*

BY J. L. WALSH

1. *Introduction.* The Green's function for a multiply connected plane region enters into many investigations, so that the nature of the equipotential curves, and in particular of their singularities, is of some importance. The writer's immediate interest in the subject, for instance, arises from the study of approximation of analytic functions by polynomials; the degree of best approximation and regions of convergence of sequences of polynomials of best approximation can be described directly and simply in terms of the equipotential curves of a certain Green's function.† For an arbitrary plane region the singular points of the equipotential curves are precisely the critical points of the Green's function, that is, the points where both first partial derivatives of Green's function vanish.

The results on the location of the critical points of the Green's function which we here prove are analogous to and proved by the use of certain well known results on the location of the roots of the derivative of a polynomial, and are not difficult to establish. However, they seem not to be known even among those versed in potential theory, and are therefore set forth here.

2. *The Green's Function.* Let R be a region of the plane, and let $P:(a, b)$ be a point of R. The Green's function $G(x, y)$ for R with pole at P is the function which is harmonic‡ interior to R except at P, continuous in the corresponding closed region except at P, zero on the boundary of R, and in some neighborhood of P can be expressed as $U(x, y)-\frac{1}{2}\log[(x-a)^2+(y-b)^2]$, where $U(x, y)$ is harmonic at P. The Green's function need not exist for an arbitrary region, but, if existent, is unique. The Green's function for a region is invariant (except for a constant factor) under conformal transformation. In particular this invariant property

* Presented to the Society, June 23, 1933.

† Walsh and Russell, forthcoming in the Transactions of this Society.

‡ A function is *harmonic at a point* if in some neighborhood of that point it is continuous together with its first and second partial derivatives and satisfies Laplace's equation. A function is *harmonic in a region* if it is harmonic at every point of that region. A *region* is an open connected point set.

gives a meaning for $G(x, y)$ when P is at infinity. The Green's function $G(x, y)$ for an infinite region R with pole P at infinity is harmonic interior to R except at infinity, continuous in the corresponding closed region except at P, zero on the boundary of R, and in the neighborhood of P can be expressed in the form $U(x, y)+\frac{1}{2}\log(x^2+y^2)$, where $U(x, y)$ is harmonic at P.

The Green's function admits the following physical interpretation. Let the boundary B of R (considered of electrically conducting material) be kept at potential zero. The potential at an arbitrary point of R due to a unit charge at P and the charge on B induced by it is $G(x, y)$. The points of equilibrium in the corresponding field of force are the critical points of $G(x, y)$. The induced charge on B is non-positive at all points of B (if this physical interpretation is justified and B is smooth), for $G(x, y)$ is everywhere positive interior to R, and approaches zero as (x, y) approaches B.

3. *Approximation by Lemniscates.* Rather than give proofs of our results based on the physical interpretation just given, which is not difficult (compare the references given below), we prefer to give more rigorous proofs based on analytic methods.

We mean by a *lemniscate* a locus of the form

$$(1) \quad |p(z)| = \mu > 0, \qquad p(z) = (z-\alpha_1)(z-\alpha_2)\cdots(z-\alpha_\sigma).$$

A lemniscate thus consists of one or more mutually exterior contours, although a finite number of exceptional points may belong to several of these contours. The points α_k lie interior to the contours. By the *interior of the lemniscate* we mean the interiors of the contours, that is, the points at which we have $|p(z)|<\mu$. *If mutually exterior arbitrary Jordan curves $C_1, C_2, \cdots, C_\nu$ are given, and arbitrarily small annular regions containing these curves, then there exists a lemniscate of form* (1) *which contains the C_k in its interior and consists of precisely ν Jordan curves lying one in each of those regions.**

If R is an arbitrary infinite region whose boundary B is finite, there exists a sequence of lemniscates L_n each interior to its

* Proved in the case $\nu=1$ by Hilbert, Göttinger Nachrichten, 1897, pp. 63–70, and in the general case by Walsh and Russell, loc. cit. Compare also Faber, Münchner Berichte, 1922, pp. 157–178, and Pólya and Szegö, Journal für Mathematik, vol. 165 (1931), pp. 4–49, with the references there given.

predecessor, each lying in R, and such that each point of R is interior to only a finite number of the L_n. It is then well known that the Green's function $G_n(x, y)$ with pole at infinity for the exterior of L_n approaches uniformly (in any closed region interior to R) the Green's function $G(x, y)$ with pole at infinity for R. It is also true, and follows from the method of proof of approximation by lemniscates, that if R is symmetric with respect to a point or line, then the lemniscates L_n can be chosen symmetric with respect to that point or line.

4. *Critical Points of the Green's Function.* For the exterior: $|p(z)| > \mu$ of the lemniscate (1), the Green's function with pole at infinity is $(1/\sigma) \log |p(z)/\mu|$, where σ is the degree of $p(z)$, as may be verified immediately. The critical points of this function are precisely the roots of $p'(z)$, concerning the location of which many theorems are known, in terms of the location of the α_k.

Let us now show that *for the region R of §3, each critical point of $G(x, y)$ in R is the limit of critical points of the $G_n(x, y)$.* Let P_1 be a critical point of $G(x, y)$ interior to R. There exists some circle C whose center is P_1 containing on or within it no point of B. All the points of L_n, for n sufficiently large, lie exterior to C, so for n sufficiently large all the functions $G_n(x, y)$ are harmonic interior to C. Let $H_n(x, y)$ be the function conjugate to $G_n(x, y)$ which vanishes at P_1; then $H_n(x, y)$, if suitably defined, is harmonic and single-valued on and within C. Let us set $F_n(z) = G_n(x, y) + iH_n(x, y)$; the sequences $\partial G_n/\partial x$, $\partial G_n/\partial y$, $H_n(x, y)$, $F_n(z)$, $F_n'(z)$ converge uniformly in an arbitrary circle C' interior to C with center at P_1, and these sequences have as their respective limits interior to C the functions $\partial G/\partial x$, $\partial G/\partial y$, $H(x, y)$, $F(z)$, $F'(z)$, where $H(x, y)$ is the function conjugate to $G(x, y)$ which vanishes at P_1 and where we have $F(z) = G(x, y) + iH(x, y)$. The function $F'(z)$ vanishes at P_1, by hypothesis. Then by a well known theorem due to Hurwitz, since $F'(z)$ does not vanish identically, the function $F_n'(z)$ vanishes in an arbitrary neighborhood of P_1, for n sufficiently large. Our proof is complete.

5. *Analog of the Gauss-Lucas Theorem.* The Gauss-Lucas theorem is: *If $p(z)$ is an arbitrary polynomial, then all roots of the derivative $p'(z)$ lie in the smallest convex region which contains the roots of $p(z)$.*

The analog which we shall prove is as follows.

THEOREM 1. *If R is an infinite region whose boundary is finite, then all the critical points of the Green's function (if existent) for R with pole at infinity lie in the smallest convex region which contains the boundary B of R.*

If all the points of B lie *interior* to a convex region K, all the lemniscates L_n (for n sufficiently large) also lie interior to K. All the roots of the polynomials defining those lemniscates lie interior to the respective L_n and hence interior to K, so (by the Gauss-Lucas theorem) all the critical points of the $G_n(x, y)$ lie interior to K and all their limit points lie in the corresponding closed region. Theorem 1 follows at once.

Another statement of Theorem 1, which applies to a more general region and more general position of the pole of the Green's function, follows directly by an inversion.

The Green's function for an arbitrary region R with pole at a point P has no critical points in any circular region which contains the point P but contains no point of the boundary of R.

By *circular region* we mean here the interior or exterior of a circle, or a half-plane.

This theorem may give fairly accurate knowledge of the location of critical points, when taken in conjunction with other known facts, such as that the set of critical points has whatever symmetry is possessed by the region R (taken together with P), and that the Green's function of a region of connectivity ρ has no more than $\rho-1$ distinct critical points. Thus, if R is a doubly connected region bounded by two circles C_1 and C_2, the unique critical point of the Green's function $G(x, y)$ for R with pole at P lies on the circle C through P orthogonal to C_1 and C_2. By the extension of Theorem 1, this critical point lies on that particular arc of C bounded by C_1 and C_2 which does not contain P.

6. *Analog of Jensen's Theorem.* The theorem of Jensen is well known:*

Let $p(z)$ be an arbitrary real polynomial (that is, with real coefficients) and let circles be drawn having as diameters the line segments joining the conjugate imaginary roots of $f(z)$. Then all the non-real roots of the derivative $p'(z)$ lie on or within these circles.

* Jensen, Acta Mathematica, vol. 36 (1912), p. 190. Walsh, Annals of Mathematics, vol. 22 (1920), pp. 128–144.

We shall prove the following analog, which also can be readily formulated for a more general region and a more general position of the pole.

THEOREM 2. *Let $G(x, y)$ be the Green's function with pole at infinity for a region R whose boundary B is finite and symmetric in a line L. Then all critical points of $G(x, y)$ not on L lie on or within the circles whose diameters are segments joining pairs of points of B symmetric with respect to L.*

It is to be noticed that these circles (which we shall call *Jensen circles*) are not merely circles whose centers lie on L and which contain all points of the boundary of R. It is essential that the circles have the segments indicated as *diameters*. The following remark is, however, useful.*

If the circle $x^2+y^2=r^2$ has on or within it a number of points symmetric in the axis of reals, then the ellipse $x^2+2y^2=2r^2$ has on or within it all the corresponding Jensen circles with respect to the axis of reals.

In order to prove Theorem 2, it is sufficient to prove that all critical points of $G(x, y)$ lie on or within a configuration K found by replacing each Jensen circle for B of radius r by a concentric circle whose radius is $r+\epsilon$, and adjoining the region $-\epsilon \leqq y \leqq \epsilon$, where $\epsilon > 0$ is arbitrary. Choose the lemniscates L_n symmetric in L. For suitably large n, all the Jensen circles for L_n lie interior to K. Indeed, a Jensen circle which corresponds to two points a_n and b_n of L_n symmetric in L such that $|a-a_n| < \epsilon/2$, $|b-b_n| < \epsilon/2$, where a and b belong to B and are symmetric in L, lies interior to K. Theorem 2 now follows by the method of §5. The following theorem reduces to Theorem 2 by an inversion.

Let $G(x, y)$ be the Green's function with pole at P for a region R whose boundary B is symmetric (anallagmatic) in a circle C which passes through P. Then every critical point of $G(x, y)$ not on C lies on a circle Q or is separated from P by a circle Q. A circle Q is any circle passing through a pair of points of B mutually inverse in C and orthogonal to the circle through P and that pair of points.

We remark, as a complement to Theorem 2 and under the hypothesis of that theorem, that *any open interval of L exterior*

* The introduction of these ellipses is again due to Jensen (loc. cit.). Formal proof of the remark is not difficult.

to all the Jensen circles and containing no point of B contains at most one critical point of $G(x, y)$; any open interval of L exterior to all the Jensen circles and containing no point of B but bounded by points of B contains precisely one critical point of $G(x, y)$. The former remark follows directly from the corresponding fact (Walsh, loc. cit.) for the polynomials which define the approximating lemniscates L_n. The latter remark follows (if L is chosen as the axis of reals) from the fact that $G(x, y)$ vanishes at both ends of the interval in question, so $\partial G/\partial x$ vanishes at some interior point P of the interval; everywhere in R on L and hence at P we have by symmetry $\partial G/\partial y = 0$, so P is a critical point of $G(x, y)$. In the two remarks just proved, the critical point of $G(x, y)$, if existent, must be simple: $|\partial^2 G/\partial x^2| + |\partial^2 G/\partial x \partial y| \neq 0$.

The following remark, also a complement to Theorem 2, is still more general but lies somewhat deeper; the proof is omitted.

Let S be a closed segment of L neither of whose end points belongs to B or lies interior to a Jensen circle for B. Let J be the configuration consisting of S and the closed interiors of all Jensen circles which intersect S. If B consists of a finite number of components (mutually exclusive closed point sets whose complements are simply connected) and if J contains precisely N of these components, then J contains precisely $N-1$, N, or $N+1$ critical points of $G(x, y)$. If S is the segment $x_1 \leqq x \leqq x_2$, $y = 0$, the number of critical points of $G(x, y)$ belonging to J is

$$
\begin{aligned}
&N-1, \quad \text{if} \quad \partial G(x_1, 0)/\partial x < 0, \ \partial G(x_2, 0)/\partial x > 0;\\
&N, \quad \text{if} \quad [\partial G(x_1, 0)/\partial x] \cdot [\partial G(x_2, 0)/\partial x] > 0;\\
&N+1, \quad \text{if} \quad \partial G(x_1, 0)/\partial x > 0, \ \partial G(x_2, 0)/\partial x < 0.
\end{aligned}
$$

In this remark, account must be taken of the multiplicities of the critical points.

7. *Analog of Walsh's Theorem.* Another theorem of interest on the roots of the derivative of a polynomial is*

Let the circle $C_1: |z-\alpha_1| = r_1$ contain on or within it m_1 roots of a polynomial $p(z)$ and let the circle $C_2: |z-\alpha_2| = r_2$ contain on or within it all the remaining roots, m_2 in number, of $p(z)$; it is immaterial whether C_1 and C_2 contain roots of $p(z)$ other than those specified. Then all roots of $p'(z)$ lie on or within C_1, C_2, and the circle

* Walsh, Transactions of this Society, vol. 22 (1921), pp. 101–116; Comptes Rendus du Congrès des Mathématiciens (Strasbourg 1920), pp. 349–352.

$$C: \left| z - \frac{m_1\alpha_2 + m_2\alpha_1}{m_1 + m_2} \right| = \frac{m_1 r_2 + m_2 r_1}{m_1 + m_2}.$$

If these circles are mutually exterior, they contain, respectively, m_1-1, m_2-1, 1 *roots of* $p'(z)$.

We shall first be concerned with the following special case.

Let $p(z)$ *be a real polynomial whose roots lie on or within two circles* $C_1: |z-\alpha| = r$ *and* $C_2: |z-\bar{\alpha}| = r$. *Then all roots of* $p'(z)$ *lie on or within* C_1, C_2, *and* $C: |z-(\alpha+\bar{\alpha})/2| = r$. *If we have* $|\alpha-\bar{\alpha}| \geqq 4r$, *then one root of* $p'(z)$ *is real (it lies on the projection of* C_1 *on the axis of reals) and the other roots lie on or within* C_1 *and* C_2.

As a matter of fact, the first part of this theorem does not follow immediately from the preceding one, for the real polynomial $p(z)$ may have an *odd* number of real roots, so that we cannot set $m_1 = m_2$. In that case it is sufficient to apply the former theorem to the polynomial $[p(z)]^2$.

The following theorem is the precise analog for critical points of the Green's function; the proof is similar to those already given and hence is omitted.

THEOREM 3. *Let* R *be an infinite region of the plane whose boundary* B *is finite and symmetric in the axis of reals, and let* $G(x, y)$ *be the Green's function for* R *with pole at infinity. If all points of* B *lie on or within the circles* $C_1: |z-\alpha| = r$ *and* $C_2: |z-\bar{\alpha}| = r$, *then all critical points of* $G(x, y)$ *lie on or within* C_1, C_2, *and* $C: |z-(\alpha+\bar{\alpha})/2| = r$. *If we have* $|\alpha-\bar{\alpha}| \geqq 4r$, *then one critical point of* $G(x, y)$ *is real (it lies on the projection of* C_1 *on the axis of reals) and the others lie on or within* C_1 *and* C_2.

8. *Extensions.* Even if the Green's function $G(x, y)$ for a given region R does not exist, it may occur that the functions $G_n(x, y)$ defined in §3 (the sequence $G_n(x, y)$ is monotonic increasing) approach a limit $G(x, y)$ in R not the infinite constant. This new function $G(x, y)$ is then called the *generalized* Green's function for R; it is independent of the particular choice of the lemniscates L_n. Theorems 1, 2, and 3, and their proofs, are clearly valid for this generalized Green's function.

Still other applications of the theorems we have mentioned exist. For instance the following theorems can be proved by the method of §7.

Let the roots of the polynomial $p(z)$ be symmetric in the origin and lie on or within the two circles $C_1: |z-\alpha| = r$ and $C_2: |z+\alpha| = r$. Then all roots of $p'(z)$ lie on or within C_1, C_2, and $C: |z| = r$. If we have $|\alpha| \geqq 2r$, then one root of $p'(z)$ lies at the origin and the other roots lie on or within C_1 and C_2.

Let R be an infinite region whose boundary B is finite and symmetric in the origin, and let $G(x, y)$ be the Green's function for R with pole at infinity. If all points of B lie on or within the circles $C_1: |z-\alpha| = r$ and $C_2: |z+\alpha| = r$, then all critical points of $G(x, y)$ lie on or within C_1, C_2, and $C: |z| = r$. If we have $|\alpha| \geqq 2r$, then one critical point of $G(x, y)$ lies at the origin and the others lie on or within C_1 and C_2.

Many other theorems on the roots of the derivative of a polynomial can also be applied in the present situation. Let us state one further application.*

Let R be an infinite region with finite boundary B and let B have three-fold symmetry about the origin (that is, let B be unchanged by a rotation of 120° about the origin). Let all points of B lie on or within the three circles $|z-\omega h| \leqq r$, where ω is a cube root of unity. Then all critical points (exterior to these three circles) of the Green's function for R with pole at infinity lie on or within the circle $|z| \leqq (r^2+hr)^{1/2}$. In particular if we have $h \geqq 3r$, then all these critical points (except for a double critical point at the origin) lie on or within the three circles $|z-\omega h| \leqq r$.

In the present note we have used results on the location of the roots of the derivative of a polynomial to establish results on the location of the critical points of the Green's function. The reciprocal process is also possible, for if a polynomial $p(z)$ is given and if $\mu > 0$ is chosen sufficiently small, the various branches of the lemniscate $|p(z)| = \mu$ lie near the roots of $p(z)$, and the roots of the derivative $p'(z)$ are the critical points of the Green's function for the exterior of this lemniscate.

HARVARD UNIVERSITY

* See Walsh, Proceedings of the National Academy of Sciences, vol. 8 (1922), pp. 139–141.

SIMPLIFICATION OF THE SET OF FOUR POSTULATES FOR BOOLEAN ALGEBRAS IN TERMS OF REJECTION*

BY B. A. BERNSTEIN

1. *Introduction.* Some time ago, I presented† a set of four postulates for Boolean algebras expressed in terms of Sheffer's "stroke" operation, or the operation of "rejection." This set, which is a reduction of Sheffer's set of five postulates for Boolean algebras,‡ uses the stroke as the only primitive idea, besides that of *class*, and retains the characteristic of Sheffer's set of defining in terms of its primitives *all* the special Boolean elements, *zero*, the *whole*, and the *negative* of an element. It is fitting that this economical set of postulates should be as simple as possible. But this is not the case. In my effort to attain economy in the number of postulates, I paid too little attention to the matter of simplicity in statement of the postulates, with the result that one of the postulates, Postulate P_4, is unnecessarily complex. It is my object now to offer a simplification of Postulate P_4.

The simplification, it will be found, will retain all the advantages possessed by the older set. It consists merely in replacing P_4 by a proposition in which the negative elements are considerably fewer, and are more symmetrically distributed, than in P_4.

In order to prove the sufficiency of the new postulates for Boolean algebras, it will of course suffice to show that my former set can be derived from them. This derivation I obtain. However, I also derive from the new postulates Sheffer's set and the well known Whitehead-Huntington set.§ Since the sufficiency of my former set is proved by showing that it yields Sheffer's set, and since the sufficiency of Sheffer's set is proved by showing that this set yields the Whitehead-Huntington set, I thought that it might be of interest to derive the latter two sets directly from my new set. These derivations, at the same time, will exhibit the workability of the new set.

* Presented to the Society, March 18, 1933.

† Transactions of this Society, vol. 17 (1916), pp. 50–52.

‡ See the Transactions of this Society, vol. 14 (1913), pp. 481–488.

§ See the Transactions of this Society, vol. 5 (1904), pp. 288–309.

I furnish for the new postulates their *complete existential theory* (in the sense of E. H. Moore), which theory will carry with it the proofs of the *consistency* and of the *independence* of the postulates. The proof-systems for this theory will be found to be of a simple arithmetic type.

2. *The New Postulates.* The new postulates have as undefined ideas the undefined ideas of the old set, namely, a *class* K and a *binary operation* $|$. The postulates are the propositions Q_1–Q_4 below. In Postulates Q_3 and Q_4, there is to be understood the condition *if the elements involved and their indicated combinations belong to* K. The supplying of this condition is essential in the consideration of the independence and of the complete existential theory of the postulates.

POSTULATE Q_1. K contains at least two distinct elements.

POSTULATE Q_2. If a, b are elements of K, $a\,|\,b$ is an element of K.

DEFINITION 1. $a' = a\,|\,a$.

POSTULATE Q_3. $a = (b\,|\,a)\,|\,(b'\,|\,a)$.

POSTULATE Q_4. $a\,|\,(b\,|\,c) = [(c'\,|\,a)\,|\,(b'\,|\,a)]'$.

3. *Sufficiency of the Postulates, Derivation of the Old Postulates.* It is seen that Postulates Q_1, Q_2, Q_3 are precisely my former postulates P_1, P_2, P_3, respectively. Hence, in order to establish the sufficiency of Postulates Q_1–Q_4 for Boolean algebras, it remains to derive from Q_1–Q_4 my former Postulate P_4, namely,

POSTULATE P_4. $a'\,|\,(b'\,|\,c) = [(b\,|\,a')\,|\,(c'\,|\,a')]'$.

This derivation I shall effect with the help of the following auxiliary theorems A and B.

THEOREM A. $a'' = a$, *where* $a'' = (a')'$.

THEOREM B. $a\,|\,b = b\,|\,a$.

The proofs of these theorems follow.

PROOF OF A. $a = (a\,|\,a)\,|\,(a'\,|\,a) = a'\,|\,(a'\,|\,a) = [(a'\,|\,a')\,|\,(a''\,|\,a')]' = a''$, by Q_3, Definition 1, Q_4, Q_3.

PROOF OF B. $a\,|\,b = a\,|\,b'' = a\,|\,(b'\,|\,b') = [(b''\,|\,a)\,|\,(b''\,|\,a)]' = [(b\,|\,a)\,|\,(b\,|\,a)]' = (b\,|\,a)'' = b\,|\,a$, by A, Definition 1, Q_4, A, Definition 1, A.

The proof of P_4 now follows.

PROOF OF P_4. $a'\,|\,(b'\,|\,c) = [(c'\,|\,a')\,|\,(b''\,|\,a')]' = [(c'\,|\,a')\,|\,(b\,|\,a')]' = [(b\,|\,a')\,|\,(c'\,|\,a')]'$, by Q_4, A, B.

Postulates Q_1–Q_4 are thus sufficient for Boolean algebras.

It can easily be verified that Postulates Q_1–Q_4 can be derived from Postulates P_1–P_4. Hence, Postulates P_1–P_4 and Postulates Q_1–Q_4 are, in fact, equivalent to one another.

I pass now to the derivation of Sheffer's postulates.

4. *Derivation of Sheffer's Postulates.* Sheffer's postulates are the propositions 1–5 below. In Postulates 3–5, the element a' is given by Definition 1 above, and in these postulates there is understood the condition *if the elements involved and their indicated combinations belong to K.*

POSTULATE 1. There are at least two distinct K-elements.

POSTULATE 2. Whenever a and b are K-elements, $a \mid b$ is a K-element.

POSTULATE 3. $(a')' = a$.

POSTULATE 4. $a \mid (b \mid b') = a'$.

POSTULATE 5. $[a \mid (b \mid c)]' = (b' \mid a) \mid (c' \mid a)$.

The derivation of these postulates from Q_1–Q_4 follow.

PROOF OF 1. By Q_1.

PROOF OF 2. By Q_2.

PROOF OF 3. By A.

PROOF OF 4. $a \mid (b \mid b') = [(b'' \mid a) \mid (b' \mid a)]' = [(b \mid a) \mid (b' \mid a)]' = a'$, by Q_4, A, Q_2.

PROOF OF 5. $[a \mid (b \mid c)]' = [(c' \mid a) \mid (b' \mid a)]'' = (c' \mid a) \mid (b' \mid a) = (b' \mid a) \mid (c' \mid a)$, by Q_4, A, B.

I proceed, finally, to the derivation of the Whitehead-Huntington postulates.

5. *Derivation of the Whitehead-Huntington Postulates.* The Whitehead-Huntington postulates leave undefined a *class* K and two *binary operations* $+$ and $\times$, and are the propositions Ia, Ib, $\cdots$, VI below. In Postulates IIIa–IVb is understood the condition *if the elements involved and their indicated combinations belong to K.* In V is understood the condition *if the elements z and u of* IIa *and* IIb *exist and are unique.*

POSTULATE Ia. $a+b$ is in K whenever a and b are in K.

POSTULATE Ib. ab is in K whenever a and b are in K.

POSTULATE IIa. There is an element z such that $a+z=a$ for every element a.

POSTULATE IIb. There is an element u such that $au=a$ for every element a.

POSTULATE IIIa. $a+b=b+a$.

POSTULATE IIIb. $ab=ba$.

POSTULATE IVa. $a+bc=(a+b)\ (a+c)$.

POSTULATE IVb. $a(b+c)=ab+ac$.

POSTULATE V. For every element a there is an element $\bar{a}$ such that $a+\bar{a}=u$ and $a\bar{a}=z$.

POSTULATE VI. There are at least two elements, a and b, in K such that $a\neq b$.

The proofs of Ia–VI follow.

DEFINITION 2. $a+b=(a\,|\,b)'$.

DEFINITION 3. $ab=a'\,|\,b'$.

PROOF OF Ia. By Definition 2, Definition 1, Q_2.

PROOF OF Ib. By Definition 3, Definition 1, Q_2.

PROOF OF IIa. The element $b'\,|\,b$, for any b, will serve as z. For, $a+(b'\,|\,b)=[a\,|\,(b'\,|\,b)]'=[(b'\,|\,a)\,|\,(b''\,|\,a)]''=a''=a$, by Definition 2, Q_4, Q_3, A.

PROOF OF IIb. The element $(b'\,|\,b)'$, for any b, will serve as u. For $a(b'\,|\,b)'=a'\,|\,(b'\,|\,b)''=a'\,|\,(b'\,|\,b)=[(b'\,|\,a')\,|\,(b''\,|\,a')]'=a''=a$, by Definition 3, A, Q_4, Q_3, A.

PROOF OF IIIa. $a+b=(a\,|\,b)'=(b\,|\,a)'=b+a$, by Definition 2, B, Definition 2.

PROOF OF IIIb. $ab=a'\,|\,b'=b'\,|\,a'=ba$, by Definition 3, B, Definition 3.

PROOF OF IVa. $a+bc=[a\,|\,(bc)]'=[a\,|\,(b'\,|\,c')]'=(c''\,|\,a)\,|\,(b''\,|\,a)]''=(c\,|\,a)\,|\,(b\,|\,a)=(a\,|\,b)\,|\,(a\,|\,c)=(a\,|\,b)''\,|\,(a\,|\,c)''=(a+b)'\,|\,(a+c)'=(a+b)\ (a+c)$, by Definition 2, Definition 3, Q_4, A, B, A, Definition 2, Definition 3.

PROOF OF IVb. $a(b+c)=a'\,|\,(b+c)'=a'\,|\,(b\,|\,c)''=a'\,|\,(b\,|\,c)=[(c'\,|\,a')\,|\,(b'\,|\,a')]'=[(a'\,|\,b')\,|\,(a'\,|\,c')]'=[(ab)\,|\,(ac)]'=ab+ac$, by Definition 3, Definition 2, A, Q_4, B, Definition 3, Definition 2.

PROOF OF V. The element a' will serve as $\bar{a}$. For, (1) $a'\,|\,a$ and $(a'\,|\,a)'$ may serve as z and u, respectively, by proof of IIa and proof of IIb; (2) $a+a'=(a\,|\,a')'=(a'\,|\,a)'$, by Definition 2, B; (3) $aa'=a'\,|\,a''=a'\,|\,a$, by Definition 3, A.

PROOF OF VI. By Q_1.

6. *The Complete Existential Theory of the Postulates.* The complete existential theory of postulates Q_1–Q_4 is given in the table below. In this table, a + sign in the ith place of the *character* of a system denotes that the system satisfies postulate Q_i, a − sign, that the system does not satisfy Q_i. In the concrete systems

of the table, the elements and the operations involved are all *arithmetic*. A symbol of the type $f(a, b) \pmod{p}$ in an arithmetic system denotes *the least positive integer (including 0) obtained from $f(a, b)$ by dropping multiples of p*. A blank for K and for $a \mid b$ indicates that there exists no system having the character concerned. The consistency of Q_1–Q_4 is given by system 1; the independence of Q_1, Q_2, Q_3, Q_4 is given, respectively, by systems 2, 3, 4, 5. The table follows.

System	Character	K	$a \mid b$
1	(++++)	0, 1	$ab+1 \pmod 2$
2	(−+++)	0	0
3	(+−++)	0, 1	0/0
4	(++−+)	0, 1	0
5	(+++−)	0, 1	b
6	(−−++)	0	0/0
7	(−+−+)	—	—
8	(−++−)	—	—
9	(+−−+)	0, 1	$0/\{ab+a+1\} \pmod 2$*
10	(+−+−)	0, 1, 2	$b+0/\{1-(a-a^2)(b-b^2)\} \pmod 3$*
11	(++−−)	0, 1	a
12	(−−−+)	—	—
13	(−−+−)	—	—
14	(−+−−)	—	—
15	(+−−−)	0, 1, 2	$a+0/\{1-(a-a^2)(b-b^2)\} \pmod 3$*
16	(−−−−)	—	—

* In system 9, Q_2 fails for $a=1$, $b=0$; Q_3 fails for $a=b=1$. In system 10, Q_2 fails for $a=b=2$; Q_4 fails for $a=b=0$, $c=1$. In system 15, Q_2 fails for $a=b=2$; Q_3 fails for $a=0$, $b=1$; Q_4 fails for $a=b=0$, $c=1$.

THE UNIVERSITY OF CALIFORNIA

ON SECTION A OF PRINCIPIA MATHEMATICA†

BY B. A. BERNSTEIN

1. *Introduction.* Section A of Whitehead and Russell's *Principia Mathematica*, which is concerned with the "theory of deduction" for "elementary" propositions, consists of "formal" (or "official") propositions and of "informal" (or "unofficial") statements. In previous papers,‡ I discussed the theory of deduction when that theory is considered as given solely by the "formal" propositions. Huntington§ has recently given sets of postulates for "the algebra of logic" (or "Boolean algebra," or "the classic logic of classes," or "the logic of classes");‖ and he obtained, in connection with these postulate-sets, two "informal" systems, consisting of "formal" and of "informal" propositions of Section A. I propose to bring out certain facts concerning Section A in view of Huntington's findings. My main results are that the "formal" Section A, as given by Huntington or by me, is derivable from the logic of classes; that this "formal" Section A is not the *whole* of the logic of classes; and that Huntington's "informal" systems are not adequate either for the classic logic of classes or for "the classic logic of propositions."¶

2. *The "Formal" Section A Derivable from the Logic of Classes.* Huntington's sixth set of postulates for Boolean algebra is expressed in terms of the following undefined ideas: a *class* K, a *subclass* T of K, a *binary operation* $+$, and a *unary operation* $'$. In this set, Postulates 6.1, 6.2, 6.3, 6.4, 6.5, 6.6, 6.7 correspond

† Presented to the Society, March 18, 1933.

‡ See especially this Bulletin, vol. 37 (1931), pp. 480–488, and vol. 38 (1932), pp. 589–593. I shall refer to these papers later as Paper 1 and Paper 2, respectively.

§ See the Transactions of this Society, vol. 35 (1933), pp. 274–304. Later references to Huntington will imply this paper.

‖ I use these terms interchangeably for the general logic of classes as developed by Boole, modified by Schröder, and formulated postulationally by Huntington in his *Sets of postulates for "the algebra of logic."*

¶ I use the term "the classic logic of propositions," or simply "the logic of propositions," for Schröder's *Aussagenkalkül* as formulated in my *Sets of postulates for the logic of propositions*, Transactions of this Society, vol. 28 (1926), pp. 472–478.

respectively to my *Principia* propositions† 1.71, 1.7, 1.1, 1.2, 1.3, 1.4, 1.6, except that my "$a=1$" is replaced by "a is in T." The sixth set of postulates thus contains *all* the (independent) primitive propositions underlying the "formal" Section A. Hence, *the "formal" Section* A, *in Huntington's T-form, is derivable from the logic of classes.*

Also, it can be verified‡ that my *Principia* postulates 1.71, 1.7, 1.1, 1.2, 1.3, 1.4, 1.6 can all be derived from Huntington's first set of postulates for Boolean algebra. Hence, *the "formal" Section* A, *in my* 1*-form, is derivable from the logic of classes.*

We thus find that, as far as the "*formal*" Section A is concerned, *there is no proposition in it which is peculiar to the logic of propositions.*

3. *The "Formal" Section* A *not the Whole of the Logic of Classes.* Besides Postulates 6.1–6.7 mentioned above, Huntington's sixth set of postulates also contains the following postulate.

POSTULATE 6.8. "If T is a subclass having the five properties (6.3–6.7) just mentioned, then we have: If $a'+b$ is in T, and $b'+a$ is in T, then $a=b$."

But Postulate 6.8 is independent of 6.1–6.7, and Postulates 6.1–6.7, we saw, constitute the "formal" Section A. Hence, *the "formal" Section* A, *in the T-form, is not the whole of the logic of classes.*

I have shown, in another connection,§ that one cannot derive from my 1-form of the "formal" Section A the proposition P following: $a' \neq a$. But proposition P is a proposition in the logic of classes. Hence, *the "formal" Section* A, *in the* 1*-form, is not the whole of the logic of classes.* We find then *that the "formal" Section* A *alone, whether in the T-form of in the* 1*-form, is not adequate for the classic logic of classes.*‖

† See Paper 1, loc. cit.

‡ By using the following Boolean propositions: (a) 0, 1 are unique; (b) $1+a=a+1=1$; (c) $a+a=a$; (d) $a'+a=1$; (e) $(a+b)'=a'b'$; (f) $(a+b)+c=a+(b+c)$; (g) $a'b+a=a+b$; (h) $a' \neq a$; (i) $1'=0$, $0'=1$.

§ In Paper 2, loc. cit.

‖ It is then to be expected that in the derivation of the "the algebra of logic" in the *Principia* (Chapter *22), the authors would use primitives not found in Section A. This, indeed, is the case. Thus, in the proof of Postulate Ia of Huntington's first set of postulates for Boolean algebra, they use, in addition to the primitive propositions of Section A, also the following primitive propositions *outside* Section A: *9.1, *9.12, *9.13, *9.14, *9.15, *12.1.

4. *Inadequacy for Classes and for Propositions of Huntington's "Informal" System* $(K, \vee, \sim, \equiv)$. Huntington (in Appendix I), constructs an "informal" system for Section A out of the *Principia* propositions $*1.71$, $*1.7$, $*4.31$, $*4.33$, $*4.25$, $*4.5$, $*4.42$, $*4.2$, $*4.21$, $*4.22$, together with certain "informal" statements in Section A regarding the sign $\equiv$. He identifies these propositions with the postulates of his fourth set, and then comes to the following conclusion: "If we accept the above mentioned informal statements as a valid part of the *Principia*, we have the following theorem.

THEOREM I. *With respect to* $(K, \vee, \sim, \equiv)$, *the informal system of the Principia is a Boolean algebra."*

Huntington, further, identifies the *Principia* proposition $*5.15$ with a certain "informal" statement, and concludes: "If we accept the informal as well as the formal statements, we have

THEOREM II. *With respect to* $(K, \vee, \sim, \equiv)$, *the informal system of the Principia is a Boolean algebra containing only two non-equivalent elements."*

Now, if we examine the propositions that determine the system $(K, \vee, \sim, \equiv)$, we find that they are all satisfied *when the class K is empty*. Hence the system $(K, \vee, \sim, \equiv)$ cannot yield such a proposition as the following.

Q. There is an element u in K such that $au = a$ for every element a in K.

But proposition Q is a proposition in the classic logic of classes and in the classic logic of propositions. Hence, *the system* $(K, \vee, \sim, \equiv)$ *is inadequate for the classic logic of classes and for the classic logic of propositions.*†

5. *Inadequacy for Classes and for Propositions of Huntington's "Informal" System* $(K, T, +, ')$. Huntington gives (in Appendix

† In view of the fact that K may be empty, "only two," in Theorem II must be taken to mean "at most two." Theorem I seems to be an oversight. In identifying the postulates of his fourth set with corresponding *Principia* propositions, Huntington apparently overlooks the "trivial preliminary postulate" 4.0, demanding that K have at least two elements. Also, in rewriting the *Principia* propositions $*4.31$, $*4.33$, $\cdots$, $*4.42$, Huntington omits the sign "$\vdash$." (In rewriting $*4.2$ he retains "$\vdash$.") The *Principia* propositions used by Huntington form, it seems to me, a system $(K, \vee, \sim, \equiv, \vdash)$, *not* $(K, \vee, \sim, \equiv)$.

II) a set of seven independent postulates, expressed in terms of K, T, $+$, $'$ (K being the class of "elementary" propositions, and T a subclass of "true" propositions in K), as basis for an "informal *Principia* system." The postulates are as follows:

POSTULATE 1. If a and b are in K, then $a+b$ is in K. [$*1.71$.]

POSTULATE 2. If a is in K, then a' is in K. [$*1.7$.]

POSTULATE 3. If a, b, etc. are in K, then $b'+(a+b)$ is in T. [$*1.3$.]

POSTULATE 4. If a, b, etc. are in K, then $(a+b)'+(b+a)$ is in T. [$*1.4$.]

POSTULATE 5. If a, b, c, etc. are in K, then $(b'+c)'+\{(a+b)'+(a+c)\}$ is in T. [$*1.6$.]

POSTULATE 6. If $a+b$ is in T, then at least one of the elements a and b is in T.

POSTULATE 7. If a' is in T, then a is not in T.

Huntington observes that the first five of the above postulates correspond "precisely to the 'formal,' and the last two to 'informal' statements in the *Principia*." After deriving from the above postulates the *Principia* propositions $*1.1$, $*1.2$, $*1.5$ (the remaining primitive propositions of Section A), he says: "Any system (K, T, $+$, $'$) which satisfies Postulate 1–7 may be called an '*informal Principia system*,' since all the propositions, both 'formal' and 'informal,' in Section A of the *Principia* are deducible from these seven postulates."

But a glance at Postulates 1–7 shows that the postulates are all satisfied *when the class K is empty*. Hence, the postulates cannot yield the proposition Q above. Hence, Huntington's "informal *Principia* system," *as given by Postulates* 1–7, *is inadequate for the classic logic of classes and for the classic logic of propositions*.†

† The inadequacy in question concerns only the "informal *Principia* system" *given by Postulates* 1–7. A two-element Boolean algebra would be a logic of propositions and would, in the sense of Huntington's definition, also be an "informal *Principia* system," though the postulates for such an algebra would not be *equivalent* to Postulates 1–7.

Huntington says (p. 302): "Hence, an 'informal *Principia* system,' as above defined, contains only two 'non-equivalent' elements." In view of the fact that K may be empty, *only two* must be taken to mean *at most two*.

6. *Remark on "Informal" Systems.* There are two apparently insurmountable difficulties that must be overcome before one can obtain a logical system which would represent the combined "formal" and "informal" Section A, and which would yield the classic logic of propositions. One difficulty is that the combined "formal" and "informal" Section A *must not contain existence propositions.* For, the *Principia* (p. xiv) says: "General propositions of the form '$(p)\cdot fp$' do occur in Section A, but '$(\exists p)\cdot fp$' does not occur." This "informal" statement does not permit in Section A such propositions of the logic of propositions as proposition Q above (§4).†

The second difficulty is that the combined "formal" and "informal" Section A *contains inconsistencies.* For instance, the *Principia* (p. 93) says: "If p is any proposition, the proposition 'not-p,' or 'p is false,' will be represented by '$\sim p$'." The *Principia* here identifies "$\sim p$," which is an "elementary" proposition, by $*1.71$, with "$\vdash\cdot\sim p$," which is a non-"elementary" proposition. Or, to put the matter a little differently, the *Principia* here identifies "p" with "$\vdash\cdot p$," and thus violates the distinction between "p" and "$\vdash\cdot p$" which it carefully made earlier (p. 92).‡

There are minor difficulties connected with the finding of a logical system equivalent to the combined "formal" and "informal" Section A. For instance, *precisely* what are *all* the "informal" statements in Section A? May (or must) "informal" statements *outside* Section A be included in the proposed system, provided such outside statements have a bearing on Section A? If so, *precisely what* are these statements?

THE UNIVERSITY OF CALIFORNIA

---

† In my Paper 1 (loc. cit.), I transcribed the *Principia* proposition $*1.1$ in the form 1.1, which is an existence proposition. In this transcription, however, I was concerned only with the *"formal"* Section A. I saw an existence notion implied in $*1.1$, and I disregarded *Principia*, p. xiv.

‡ Huntington reconciles "$a\supset b$" the *element* with "$a\supset b$" the *relation*, "which are not in precise agreement," by the following definitions.

DEFINITION. $a\supset b$ means $a'+b$;

DEFINITION. (a implies b) means ($a'+b$ is in T).

It seems to me that these definitions merely *recognize* the disagreement. The definitions merely say that "$a\supset b$" is an *element*, namely $a'+b$, and that "a implies b" is a *relation*, namely "$\vdash\cdot a'+b$."

ON NATURAL FAMILIES OF CURVES

BY C. H. ROWE

1. *Introduction.* A *natural family* of curves in a Riemannian space V_N of N dimensions is a family which consists of the ∞^{2N-2} extremals of an integral $\int \mu ds$, where μ is a function of position and ds is the element of length. If a system of ∞^M curves is given in V_N, where $M \leqq 2N-2$, there does not in general exist a natural family to which all these curves belong; and the present paper is concerned with the problem of finding conditions for the existence of such a family. This problem is equivalent, as will be seen, to that of finding conditions for the possibility of representing V_N conformally on a second Riemannian space so that the curves of the second space that correspond to the curves of the given system are geodesics.

In the case where $M=2N-2$ and $N>2$, a condition is given by the extensions to Riemannian space of the so-called* theorem of Thompson and Tait and the converse theorem of Kasner. According to these theorems, a system of ∞^{2N-2} curves in V_N $(N>2)$ is a natural family if and only if the ∞^{N-1} curves of the system that cut an arbitrary V_{N-1} normally are the orthogonal trajectories of a family of a single infinity of V_{N-1}'s, or in other words, form a normal congruence.† This result is applicable only when the given system contains ∞^{2N-2} curves; and it is not valid for $N=2$, since the condition is satisfied by an arbitrary system of ∞^2 curves on a surface. The result that we shall obtain in what follows is not subject to these limitations, and it yields immediately the theorem of Thompson and Tait in the cases where this is applicable.

Our argument will be mainly synthetic in form; and we shall make no attempt at a rigorous discussion of the minimum assumptions under which our results hold, contenting ourselves with supposing that the functions that we introduce, explicitly

* J. A. Schouten (Nieuw Archief, (2), vol. 15 (1928), p. 97) points out that this theorem was first given by Lipschitz.

† See E. Kasner, Transactions of this Society, vol. 11 (1910), p. 121, J. A. Schouten, loc. cit., and W. Blaschke, Nieuw Archief, (2), vol. 15 (1928), p. 202. The methods of the present paper are similar to those used by Blaschke.

or implicitly, have whatever degree of regularity is required in order to justify our reasoning, and restricting ourselves whenever it is necessary to a sufficiently limited region of V_N.

2. *Conditions for Belonging to a Natural Family.* When we consider a system (S) of ∞^M curves, we shall suppose that at least one curve of (S) passes through each point of V_N, thus excluding the cases in which all the curves of (S) lie in a sub-space; and we shall suppose that at most one curve of (S) passes through a given point in a given direction. We thus have

$$N - 1 \leqq M \leqq 2N - 2.$$

We shall further assume (when $M > N-1$) that, if two curves of (S) passing through a point are given, it is possible for a variable curve of (S) which passes through this point to move continuously from coincidence with one of these curves into coincidence with the other. When $M = N-1$, the system of curves is called a congruence; and we shall then assume that a unique curve of (S) passes through each point. When $M = 2N-2$, we shall suppose that a unique curve of (S) passes through each point in each direction.

Suppose for the present that $N > 2$. By a *surface of the system* (S) we shall mean a surface (or V_2) that can be generated by a variable curve of (S) which moves with one degree of freedom; and we shall call the ∞^1 positions of the variable curve the *generators* of the surface.* If the initial and final positions of the generating curve coincide, we shall call the surface a *tube.* We shall have to consider the curves on a surface of (S) that cut the generators orthogonally, and we shall call these the *orthogonal curves* of the surface.

When a surface of (S) is a tube, it may happen that the orthogonal curve that starts from a point P returns to P after encircling the tube, and is thus a closed curve; but in general this will not happen. Instead, if an orthogonal curve is prolonged indefinitely in either sense, it will encircle the tube continually, never intersecting itself, in a manner not unlike that in which a circular helix encircles the circular cylinder on which it lies.

* If a surface can be generated in more than one way by curves of (S), it will always be clear that a particular mode of generation is being considered.

The theorem that we wish to prove may now be stated as follows.

In order that there should exist a natural family in V_N to which the curves of a given system belong, it is necessary and sufficient that, whenever one orthogonal curve on a tube of the system is closed, all the remaining orthogonal curves on that tube should also be closed.

This theorem holds without essential change when $N=2$, although it has been expressed in language appropriate to three or more dimensions. In the case where $M=N=2$, the theorem may be stated, perhaps more suitably, by saying that a system of ∞^2 curves on a surface is a natural family if, and only if, the ∞^1 curves of the system that cut an arbitrary closed curve normally are the orthogonal trajectories of a family of ∞^1 closed curves.

The case where $N=2$, $M=1$ may be conveniently taken with the case where, for $N>2$, the system is a normal congruence. In the former case we may regard the condition of our theorem as being fulfilled. In the latter case it is always fulfilled, since every tube of a normal congruence has closed orthogonal curves. In each of these cases, as we shall see, it is possible to find a natural family containing the given system, but, in contrast to the remaining cases, this family is not unique.

That our theorem yields the theorem of Thompson and Tait follows from the fact that all the tubes belonging to a congruence have closed orthogonal curves when the congruence is normal, and in no other cases.* Suppose that a system (S) of ∞^{2N-2} curves satisfies our condition, and consider the congruence formed by the curves of (S) that cut a given V_{N-1} normally. On any tube of this congruence, one, and therefore every, orthogonal curve is closed; and consequently the congruence is normal. Conversely, supposing that the system (S) satisfies the con-

* Consider a congruence (C), and suppose that every orthogonal curve on every tube of (C) is closed. On a fixed curve C_0 of (C) take a fixed point P_0, and let C be a variable curve of (C). On any surface of (C) of which C_0 and C are generators draw the orthogonal curve that passes through P_0, and let it meet C in P. The point P does not depend on our choice of this surface, for the tube formed by two such surfaces has closed orthogonal curves. As C varies in (C), P describes a V_{N-1} which clearly cuts all the curves of (C) normally. If we vary P_0 along C_0, we get a family of a single infinity of V_{N-1}'s which cut the curves of (C) normally.

dition of Thompson and Tait, consider a tube of (S) on which one orthogonal curve is closed. If we draw a V_{N-1} through this curve cutting the generators of the tube normally, the tube belongs to the congruence formed by the curves of (S) that cut this V_{N-1} normally; and since this congruence is normal, all the orthogonal curves of the tube are closed.

3. *Proof of Necessity.* The necessity of the condition of our theorem is easily proved. Let the metric of V_N be defined in any system of coordinates x^i by the formula

$$ds^2 = g_{ij}dx^i dx^j,$$

and consider the natural family (S) formed by the extremals of $\int \mu ds$, where μ is a function of the variables x^i, which we shall suppose not to vanish or become infinite in the region under consideration. Consider a second Riemannian space V_N', in which the element of length ds' is given by

$$ds'^2 = \mu^2 g_{ij}dx^i dx^j = \mu^2 ds^2.$$

The correspondence between these two spaces in which corresponding points have the same coordinates is conformal. It is therefore sufficient to prove that the system (S') in V_N' that corresponds to (S) satisfies the condition of our theorem. Now the curves of (S'), being extremals of $\int ds'$, are geodesics of V_N'; and the presence of one closed orthogonal curve on a tube generated by geodesics implies that all the orthogonal curves of this tube are closed. This is an immediate consequence of the theorem of Gauss that two orthogonal trajectories of a system of ∞^1 geodesics on a surface intercept equal arcs on all these geodesics.

4. *A Lemma.* Before considering the sufficiency of our condition, we must establish a preliminary result. We shall show that, if (S) is a system of curves which satisfies our condition, and if P_0 and P are two given points, it is possible to find a surface of (S) which has an orthogonal curve passing through P_0 and P, except in the case where (S) is a normal congruence and in the case where $N=2$, $M=1$.

Since this result is immediate when $N=2$, $M=2$, we shall suppose that $N>2$. Consider first the case where (S) is a congruence which is not normal. Let S_0 and S be the curves of (S) that pass through P_0 and P respectively, and let Σ be any sur-

face of (S) of which these curves are generators. Let the orthogonal curve on Σ that passes through P_0 meet S in a point Q, which we may suppose not to coincide with P. We can find a tube of (S) which contains the curve S as a generator, and which has no closed orthogonal curves; for if not, the orthogonal curves of every tube of which S is a generator would be closed, and (S) would therefore be a normal congruence. Let T be such a tube, and let the orthogonal curve P_0Q that we are considering on Σ be prolonged beyond the point Q by allowing it to encircle the tube T a certain number of times, until it cuts S in a point Q' which lies on the side of P remote from Q. This is clearly possible, if we allow the curve to encircle T a sufficient number of times in the right sense; for if not, the curve would cut S in an unending sequence of points lying between P and Q. These points would have a limiting point, and, for reasons of continuity, the orthogonal curve of T through this limiting point would be closed. Supposing that P lies between Q and Q', we shall allow the tube T to vary continuously, always passing through S, and to contract until it reduces to the curve S. While this happens, the point Q' moves continuously along S, and ultimately coincides with Q. At some stage in this process Q' must pass through P; and at this stage we have a tube which has an orthogonal curve joining Q to P. The surface made up of the portion of Σ between the curves S_0 and S and of the tube T has thus an orthogonal curve P_0QP which passes through P_0 and P, as is required.

Suppose now that (S) contains ∞^M curves, where $M>N-1$. It will easily be seen that we can form with curves of (S) a congruence (C) which is not normal, and which, like (S), satisfies the condition of our theorem. We can then find a surface belonging to (C), and therefore to (S), which has an orthogonal curve joining the two given points.

If, however, the system (S) is a normal congruence, or if $N=2$, $M=1$, it will not be possible to find a surface of (S) satisfying our requirement, unless one of the V_{N-1}'s that cut the curves of (S) normally passes through the two given points.

5. *Proof of Sufficiency.* We shall suppose that on any tube of a system (S) either all or none of the orthogonal curves are closed, and we shall show that we can define a function μ of position throughout V_N such that the curves of (S) belong to the

natural family formed by the extremals of $\int \mu ds$. We shall state the proof for $N>2$, without troubling to indicate the merely verbal changes that would be desirable if N were equal to 2.

We shall suppose firstly that (S) is not a normal congruence or a system of ∞^1 curves on a surface. Take a fixed point P_0, and choose an arbitrary positive value $\mu(P_0)$ for the function μ at P_0. Let P be any point, and find a surface Σ of (S) which has an orthogonal curve joining P_0 to P. Consider on Σ a second orthogonal curve, which we shall allow to tend to coincide with the first. Let the lengths of the arcs that these curves intercept on the generators of Σ that pass through P_0 and P be δs_0 and δs. We may assume that, as the second orthogonal curve tends to the first, the ratio $\delta s_0/\delta s$ tends to a limit; and we shall show that this limit depends only on the point P (the point P_0 being fixed), and not on our choice of the surface Σ. We must thus show that this limit is unaltered if we replace Σ by another surface Σ' which also has an orthogonal curve joining P_0 to P. If (S) is a congruence, the surfaces Σ and Σ', having in common the curves of the congruence that pass through P_0 and P, form a tube; and since the orthogonal curves of this tube are closed, the truth of our assertion is clear. In the general case $(M>N-1)$ we shall construct a tube of which Σ and Σ' form part. We connect the generators of Σ and Σ' that pass through P_0 by a surface of (S) whose generators all pass through P_0; and, similarly, we connect the generators of Σ and Σ' that pass through P by a surface of (S) whose generators all pass through P. These two new surfaces, together with Σ and Σ', form a tube of (S) on which one, and therefore every, orthogonal curve is closed. Let the second orthogonal curve that we have taken on Σ be continued around this tube until it closes. We thus get a second orthogonal curve on Σ'. If $\delta s_0'$ and $\delta s'$ are the lengths of the arcs that the two orthogonal curves on Σ' intercept on the generators of Σ' through P_0 and P, we have to show that

$$\lim \frac{\delta s_0'}{\delta s'} = \lim \frac{\delta s_0}{\delta s}.$$

We notice that, if a surface is generated by curves passing through a fixed point, the arcs intercepted on two of these curves between the fixed point and an orthogonal trajectory of the

curves have a ratio which tends to unity as the lengths of the arcs tend to zero. We thus have the equations

$$\lim \frac{\delta s_0'}{\delta s_0} = 1, \quad \lim \frac{\delta s'}{\delta s} = 1,$$

from which the required conclusion follows at once.

Since $\lim \delta s_0/\delta s$ depends only on the point P, we may define the value $\mu(P)$ of the function μ at any point P by the equation

$$\mu(P) = \mu(P_0) \lim \frac{\delta s_0}{\delta s},$$

which may also be written in the equivalent form

$$\mu(P)ds = \mu(P_0)ds_0.$$

It is clear from this that, if we take any surface of (S), and consider the arcs intercepted on two of its generators by two of its orthogonal curves, the values of $\int \mu ds$ arising from these two arcs are equal. We shall suppose that the function μ does not vanish or become infinite in the region that we are considering, and we shall introduce, as we did in §3, a second Riemannian space V_N' in conformal correspondence with V_N, the element of length ds' in V_N' being given by $ds' = \mu ds$. The system (S') in V_N' that corresponds to (S) has the property that the arcs intercepted on the generators of a surface of (S') by two of its orthogonal curves are all equal. In virtue of the converse of the theorem of Gauss that we quoted in §3, every curve of (S') is therefore a geodesic of any surface of (S') of which it is a generator. This implies that the first curvature-vector of a curve of (S') either vanishes or is normal to every surface of (S') of which the curve is a generator. The latter alternative is clearly impossible, and therefore every curve of (S') is a geodesic of V_N'. It follows that the curves of (S) belong to the natural family formed by the extremals of $\int \mu ds$.

Since we have defined the function μ uniquely except for a constant factor, our procedure leads to a unique natural family. It will be seen that this is in fact the only natural family to which all the curves of (S) belong.

It remains to consider the case where (S) is a normal congruence, and the case where $N=2$, $M=1$. In these cases the

condition of our theorem is always verified, as we have already remarked, and we shall show that we can find a natural family which contains the given curves; but this family is no longer unique. If we try to use the same method as before, we find that we can define the value of μ only at points of the normal V_{N-1} that passes through P_0. Instead, therefore, we take an arbitrary curve which meets each normal V_{N-1} once, and we define the function μ arbitrarily along this curve. We then use our former method to define the value of μ at any point P in terms of the assigned value of μ at the point where the normal V_{N-1} that passes through P meets this curve. Reasoning as before, we can then prove that the curves of (S) are extremals of $\int \mu ds$. We can thus find a natural family depending on an arbitrary function of one variable which contains the curves of the given system.*

We may remark, in conclusion, that a modification of the methods that we have used will allow us to consider a more general problem. If we are given a system (S) of curves in V_N and a transversality relation, we may ask whether there is a first order problem of the calculus of variations such that the family of its ∞^{2N-2} extremals contains the curves of (S), and such that its transversality relation coincides with the given one. If we replace the relation of orthogonality by the given transversality relation, we are led to consider on any surface of (S) the curves that are cut transversally by the generators, and we may call these the transversal curves of the surface. It is then possible to generalize the theorem of Kneser on transversals and its converse† in the same way as our previous result gener-

* It is perhaps worth noticing that, in the cases at present under consideration, a natural family containing the given system (S) is uniquely determined if we are given one curve Γ which belongs to the natural family, but not to (S), and which cuts each normal V_{N-1} once, but not normally. We recall the fact that the first curvature-vector of an extremal of $\int \mu ds$ is the component normal to the curve of the gradient of $\log \mu$. This gradient is determined at each point of Γ, because we know its components normal to each of two distinct directions, namely, that of Γ and that of the curve of (S) through the point. The value of μ is therefore determined along Γ except for a constant factor, and consequently the natural family is determined uniquely. It will be seen that, when $N>2$, the curve Γ may not be given arbitrarily.

† See J. Douglas, Transactions of this Society, vol. 29 (1927), p. 401, and W. Blaschke, loc. cit.

alizes the theorem of Thompson and Tait. We can prove, in fact, that a condition for an affirmative answer to our question is that, on any tube of (S), either all or none of the transversal curves should be closed.

TRINITY COLLEGE,
DUBLIN, IRELAND

ON THE CONDITION THAT TWO ZEHFUSS MATRICES BE EQUAL

BY D. E. RUTHERFORD

1. *Introduction.* In a recent paper* Williamson has considered matrices whose sth compounds are equal. The present paper considers the somewhat analogous problem of finding the conditions that two Zehfuss matrices be equal.

Suppose that R is a matrix of n_1 rows and m_1 columns whose ijth element is r_{ij}, and that P is another matrix of n_2 rows and m_2 columns. Now, if the matrix Q of n_1n_2 rows and m_1m_2 columns can be partitioned into submatrices each of n_2 rows and m_2 columns such that the ijth submatrix is $r_{ij}P$, then Q is a *Zehfuss matrix*† or the *direct product matrix*‡ of R and P. We shall write

$$Q = R\langle P\rangle = \langle P\rangle R.$$

In general, however, $R\langle P\rangle \neq \langle P\rangle R$.

It is the purpose of this paper to find out under what conditions the matrix equation

$$A\langle B\rangle = C\langle D\rangle$$

is true. That is, we shall find the most general form of the matrices A, B, C, D when the above equation holds.

2. *The Simplest Case.* We shall begin by considering the simplest case, where A, B, C, D are row vectors, where A and D are of order m_1, where B and C are of order m_2, and where

$$(m_1, m_2) = 1;$$

that is to say, m_1 and m_2 are prime to one another. Suppose that

---

* J. Williamson, this Bulletin, vol. 39 (1933), p. 109.

† G. Zehfuss, Zeitschrift für Mathematik und Physik, vol. 3 (1858), p. 298.

‡ L. E. Dickson, *Algebras and Their Arithmetics*, p. 119.

$$A = [a_1, a_2, \cdots, a_{m_1}], \quad B = [b_1, b_2, \cdots, b_{m_2}],$$
$$C = [c_1, c_2, \cdots, c_{m_2}], \quad D = [d_1, d_2, \cdots, d_{m_1}],$$

and let $m_2 > m_1$. From definition, we have

$$A\langle B\rangle = [a_1b_1, a_1b_2, \cdots, a_1b_{m_2}, a_2b_1, \cdots, a_2b_{m_2}, \cdots, a_{m_1}b_1, \cdots, a_{m_1}b_{m_2}]$$

and

$$C\langle D\rangle = [c_1d_1, c_1d_2, \cdots, c_1d_{m_1}, c_2d_1, \cdots, c_2d_{m_1}, \cdots, c_{m_2}d_1, \cdots, c_{m_2}d_{m_1}].$$

Identifying these two row vectors element by element, we get m_1m_2 equations, determining the relations which must hold between the elements of A, B, C, D. Amongst these m_1m_2 equations, consider the following:

$$\begin{aligned} \alpha_1b_1 &= \gamma_1d_1, \\ \alpha_2b_2 &= \gamma_2d_1, \\ &\cdots\cdots\cdots\cdots\cdots \\ \alpha_{m_2-m_1+1}b_{m_2-m_1+1} &= \gamma_{m_2-m_1+1}d_1, \end{aligned} \tag{1}$$

where each α represents some one of $a_1, \cdots, a_{m_1}$, and each γ some one of $c_1, \cdots, c_{m_2}$. A little consideration will show that no two α's represent the same a and no two γ's represent the same c. It is obvious that $\alpha_1 = a_1$ and $\gamma_1 = c_1$. From equations (1) and from the construction of $A\langle B\rangle$ and $C\langle D\rangle$ it follows that

$$\begin{aligned} \alpha_1[b_1, \cdots, b_{m_1}] &= \gamma_1[d_1, \cdots, d_{m_1}], \\ \alpha_2[b_2, \cdots, b_{m_1+1}] &= \gamma_2[d_1, \cdots, d_{m_1}], \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ \alpha_{m_2-m_1+1}[b_{m_2-m_1+1}, \cdots, b_{m_2}] &= \gamma_{m_2-m_1+1}[d_1, \cdots, d_{m_1}]. \end{aligned} \tag{2}$$

From equations (2) we deduce

$$b_2 = \frac{\gamma_1}{\alpha_1}d_2 = \frac{\gamma_2}{\alpha_2}d_1, \quad \text{whence} \quad \frac{d_2}{d_1} = \frac{\gamma_2\alpha_1}{\gamma_1\alpha_2} = s, \text{ say,}$$

and

$$b_3 = \frac{\gamma_2}{\alpha_2}d_2 = \frac{\gamma_3}{\alpha_3}d_1, \quad \text{whence} \quad \frac{d_2}{d_1} = \frac{\gamma_3\alpha_2}{\gamma_2\alpha_3} = s,$$

and so on. In this way we can show that

$$(3) \qquad s = \frac{\gamma_2\alpha_1}{\gamma_1\alpha_2} = \frac{\gamma_3\alpha_2}{\gamma_2\alpha_3} = \cdots = \frac{\gamma_{m_2-m_1+1}\alpha_{m_2-m_1}}{\gamma_{m_2-m_1}\alpha_{m_2-m_1+1}}.$$

Again, from equations (2), we find that

$$b_3 = \frac{\gamma_1}{\alpha_1}d_3 = \frac{\gamma_2}{\alpha_2}d_2, \quad \text{whence} \quad \frac{d_3}{d_2} = \frac{\gamma_2\alpha_1}{\gamma_1\alpha_2} = s.$$

By a repetition of such an argument, we can show that

$$s = \frac{d_2}{d_1} = \frac{d_3}{d_2} = \cdots = \frac{d_{m_1}}{d_{m_1-1}}.$$

Hence

$$D = d_1[1, s, s^2, \cdots, s^{m_1-1}].$$

Equating the last elements in each of the matrix equations (2), we find that by equations (3), since $d_{m_1} = d_1 s^{m_1-1}$,

$$b_{m_1+1} = \frac{\gamma_2}{\alpha_2}d_{m_1} = \frac{\gamma_2\alpha_1}{\gamma_1\alpha_2}\frac{\gamma_1}{\alpha_1}d_1 s^{m_1-1} = \frac{\gamma_1}{\alpha_1}d_1 s^{m_1} = \frac{c_1}{a_1}d_1 s^{m_1},$$

$$b_{m_1+2} = \frac{\gamma_3}{\alpha_3}d_{m_1} = \frac{\gamma_3\alpha_2}{\gamma_2\alpha_3}\frac{\gamma_2\alpha_1}{\gamma_1\alpha_2}\frac{\gamma_1}{\alpha_1}d_1 s^{m_1-1} = \frac{\gamma_1}{\alpha_1}d_1 s^{m_1+1} = \frac{c_1}{a_1}d_1 s^{m_1+1},$$

. . . . . .

$$b_{m_2} = \frac{\gamma_{m_2-m_1+1}}{\alpha_{m_2-m_1+1}}d_{m_1} = \cdots = \frac{\gamma_1}{\alpha_1}d_1 s^{m_2-1} = \frac{c_1}{a_1}d_1 s^{m_2-1}.$$

Hence, since

$$[b_1, \cdots, b_{m_1}] = \frac{\gamma_1}{\alpha_1}d_1[1, s, \cdots, s^{m_1-1}] = \frac{c_1}{a_1}d_1[1, s, \cdots, s^{m_1-1}],$$

we may write

$$B = [b_1, \cdots, b_{m_2}] = \frac{c_1}{a_1}d_1[1, s, \cdots, s^{m_2-1}] = b_1[1, s, \cdots, s^{m_2-1}].$$

We have now shown that with the possible exceptions of the elements $a_2b_1, a_3b_1, \cdots, a_{m_1}b_1$, every element of $A\langle B\rangle$ is s times the preceding element; that is, every element, with the possible exception of the

$$(4) \qquad (m_2+1)\text{th}, \quad (2m_2+1)\text{th}, \cdots, \quad (m_2(m_1-1)+1)\text{th}$$

elements. But this is also true of $C\langle D\rangle$ with the possible exception of the

$$(5) \qquad (m_1+1)\text{th}, \quad (2m_1+1)\text{th}, \cdots, (m_1(m_2-1)+1)\text{th}$$

elements. But, since $(m_1, m_2)=1$, no member of the set (4) is a member of the set (5); and since the elements of $A\langle B\rangle$ are identical with those of $C\langle D\rangle$, there are no exceptions and every element of $A\langle B\rangle$ or $C\langle D\rangle$ is s times the preceding element.

Hence

$$a_2b_1 = sa_1b_{m_2} = sa_1b_1s^{m_2-1},$$

so that

$$a_2 = a_1s^{m_2},$$

and similarly

$$a_3 = a_2s^{m_2} = a_1s^{2m_2},$$
$$\cdots\cdots\cdots\cdots$$
$$a_{m_1} = a_1s^{m_2(m_1-1)}.$$

It follows that

$$A = a_1[1, s^{m_2}, s^{2m_2}, \cdots, s^{m_2(m_1-1)}],$$

and similarly

$$C = c_1[1, s^{m_1}, s^{2m_1}, \cdots, s^{m_1(m_2-1)}].$$

We have not yet considered all the m_1m_2 equations connecting the elements of A, B, C, D; but, since the above values of A, B, C, D give a solution for any values of the arbitrary quantities a_1, c_1, d_1, s, they also give the most general solution.

3. *A More General Case.* We shall now consider a more general case, where A and D are rectangular matrices of n_1 rows and m_1 columns, where B and C are rectangular matrices of n_2 rows and m_2 columns, and where $(n_1, n_2)=1$, and $(m_1, m_2)=1$. Let

$$A = \begin{bmatrix} a_{11}, & a_{12}, \cdots, a_{1m_1} \\ \cdots\cdots\cdots & \\ a_{n_11}, & a_{n_12}, \cdots, a_{n_1m_1} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11}, & b_{12}, \cdots, b_{1m_2} \\ \cdots\cdots\cdots & \\ b_{n_21}, & b_{n_22}, \cdots, b_{n_2m_2} \end{bmatrix},$$

$$C = \begin{bmatrix} c_{11}, & c_{12}, \cdots, c_{1m_2} \\ \cdots\cdots\cdots & \\ c_{n_21}, & c_{n_22}, \cdots, c_{n_2m_2} \end{bmatrix}, \quad D = \begin{bmatrix} d_{11}, & d_{12}, \cdots, d_{1m_1} \\ \cdots\cdots\cdots & \\ d_{n_11}, & d_{n_12}, \cdots, d_{n_1m_1} \end{bmatrix}.$$

By equating the first rows of $A\langle B\rangle$ and $C\langle D\rangle$ we obtain from §2 the relation

$$[b_{11}, b_{12}, \cdots, b_{1m_2}] = b_{11}[1, s, \cdots, s^{m_3-1}]. \tag{6}$$

Similarly, by equating the second rows of $A\langle B\rangle$ and $C\langle D\rangle$ we deduce that

$$[b_{21}, b_{22}, \cdots, b_{2m_2}] = b_{21}[1, s, \cdots, s^{m_3-1}], \tag{7}$$

for the a's and c's occurring are the same in both rows and hence the s must be the same in (7) as in (6). Proceeding in this way with the rows of $A\langle B\rangle$ and $C\langle D\rangle$, we obtain eventually

$$B = \begin{bmatrix} b_{11}, & b_{12}, \cdots, b_{1m_2} \\ \cdots\cdots\cdots \\ b_{n_21}, & b_{n_22}, \cdots, b_{n_2m_2} \end{bmatrix} = \begin{bmatrix} b_{11} \\ b_{21} \\ \cdot \\ \cdot \\ b_{n_21} \end{bmatrix} [1, s, \cdots, s^{m_3-1}].$$

We shall find it more convenient to denote the first factor on the right hand side by $\{b_{11}, b_{21}, \cdots, b_{n_21}\}$, as is frequently done, that is, the curly brackets denote a column vector.

Now, by equating the first columns in $A\langle B\rangle$ and $C\langle D\rangle$, we obtain, in the same manner as (6) was obtained, the relation

$$\{b_{11}, b_{21}, \cdots, b_{n_21}\} = b_{11}\{1, t, t^2, \cdots, t^{n_3-1}\},$$

where t is a new arbitrary quantity. Hence

$$B = b_{11}\{1, t, \cdots, t^{n_3-1}\}[1, s, \cdots, s^{m_3-1}],$$

and in the same manner

$$D = d_{11}\{1, t, \cdots, t^{n_1-1}\}[1, s, \cdots, s^{m_1-1}],$$

$$A = a_{11}\{1, t^{n_2}, t^{2n_2}, \cdots, t^{n_2(n_1-1)}\}[1, s^{m_2}, s^{2m_2}, \cdots, s^{m_2(m_1-1)}],$$

$$C = c_{11}\{1, t^{n_1}, t^{2n_1}, \cdots, t^{n_1(n_2-1)}\}[1, s^{m_1}, s^{2m_1}, \cdots, s^{m_1(m_2-1)}],$$

where $a_{11}b_{11} = c_{11}d_{11}$. It follows that since the above values of A, B, C, D give a solution of $A\langle B\rangle = C\langle D\rangle$, for any values of a_{11}, c_{11}, d_{11}, s, t, they give the most general solution.

4. *The Most General Case.* We shall now consider the most general case and show that its solution is dependent upon the one just obtained. Suppose that the matrices A, B, C, D have

n_1, n_2, n_3, n_4 rows and m_1, m_2, m_3, m_4 columns respectively. Since $A\langle B\rangle = C\langle D\rangle$,

$$n_1 n_2 = n_3 n_4, \text{ and } m_1 m_2 = m_3 m_4. \tag{8}$$

Let the highest common factor of n_1 and n_3 be k_1. We write this $(n_1, n_3) = k_1$. Let $n_1 = \nu_1 k_1$ and $n_3 = \nu_3 k_1$, where $(\nu_1, \nu_3) = 1$. Similarly, let $n_2 = \nu_2 k_2$ and $n_4 = \nu_4 k_2$, where $(\nu_2, \nu_4) = 1$; let $m_1 = \mu_1 h_1$ and $m_3 = \mu_3 h_1$, where $(\mu_1, \mu_3) = 1$; and let $m_2 = \mu_2 h_2$ and $m_4 = \mu_4 h_2$, where $(\mu_2, \mu_4) = 1$. From equation (8) $\nu_1\nu_2 k_1 k_2 = \nu_3\nu_4 k_1 k_2$ and therefore $\nu_1\nu_2 = \nu_3\nu_4$. Now since $(\nu_1, \nu_3) = 1$, ν_1 must be a factor of ν_4, and since $(\nu_2, \nu_4) = 1$, ν_4 must be a factor of ν_1. Hence $\nu_1 = \nu_4$ and $\nu_2 = \nu_3$, also $(\nu_1, \nu_2) = 1$. Similarly $\mu_1 = \mu_4$ and $\mu_2 = \mu_3$, also $(\mu_1, \mu_2) = 1$. The procedure is now quite simple, although it is somewhat difficult to explain in writing. Consider the very simple case

$$[a_1, a_2, a_3, a_4, a_5, a_6]\,\langle[b_1, b_2, b_3, b_4]\rangle = [c_1, c_2, c_3, c_4]\langle[d_1, d_2, d_3, d_4, d_5, d_6]\rangle.$$

In this example, $h_1 = 2$, and we see that the above equation can be split up into the two equations

$$[a_1, a_2, a_3]\langle[b_1, b_2, b_3, b_4]\rangle = [c_1, c_2]\langle[d_1, d_2, d_3, d_4, d_5, d_6]\rangle,$$
$$[a_4, a_5, a_6]\langle[b_1, b_2, b_3, b_4]\rangle = [c_3, c_4]\langle[d_1, d_2, d_3, d_4, d_5, d_6]\rangle.$$

In this example $h_2 = 2$, and we can split up each of the above into two equations and so we can reduce this case to the following four examples of the case considered in §2:

$$[a_1, a_2, a_3]\langle[b_1, b_3]\rangle = [c_1, c_2]\langle[d_1, d_3, d_5]\rangle,$$
$$[a_1, a_2, a_3]\langle[b_2, b_4]\rangle = [c_1, c_2]\langle[d_2, d_4, d_6]\rangle,$$
$$[a_4, a_5, a_6]\langle[b_1, b_3]\rangle = [c_3, c_4]\langle[d_1, d_3, d_5]\rangle,$$
$$[a_4, a_5, a_6]\langle[b_2, b_4]\rangle = [c_3, c_4]\langle[d_2, d_4, d_6]\rangle.$$

In the most general case, we can split up the equation

$$A\langle B\rangle = C\langle D\rangle$$

into $k_1 k_2 h_1 h_2$ equations

$$A_{xy}\langle B_{zu}\rangle = C_{xy}\langle D_{zu}\rangle,$$

where $x = 1, 2, \cdots, k_1$; $y = 1, 2, \cdots, h_1$; $z = 1, 2, \cdots, k_2$; $u = 1, 2, \cdots, h_2$, and where A_{xy} is the matrix of ν_1 rows and μ_1

columns whose ijth element is $a_{(x-1)\nu_1+i,(y-1)\mu_1+j}$. That is to say, $A_{xy}=[a_{(x-1)\nu_1+i,(y-1)\mu_1+j}]$ has ν_1 rows and μ_1 columns. Similarly, $C_{xy}=[c_{(x-1)\nu_2+i,(y-1)\mu_2+j}]$ has ν_2 rows and μ_2 columns; while $B_{zu}=[b_{(i-1)k_2+z,(j-1)h_2+u}]$ has ν_2 rows and μ_2 columns and $D_{zu}=[d_{(i-1)k_2+z,(j-1)h_2+u}]$ has ν_1 rows and μ_1 columns. But, since $(\nu_1, \nu_2)=1$ and $(\mu_1, \mu_2)=1$, the most general case is composed of $k_1k_2h_1h_2$ examples of the case treated in §3. For brevity, let us write

$$E = \{1, t^{\nu_2}, \cdots, t^{\nu_2(\nu_1-1)}\}[1, s^{\mu_2}, \cdots, s^{\mu_2(\mu_1-1)}],$$

$$F = \{1, t, \cdots, t^{\nu_2-1}\}[1, s, \cdots, s^{\mu_2-1}],$$

$$G = \{1, t^{\nu_1}, \cdots, t^{\nu_1(\nu_2-1)}\}[1, s^{\mu_1}, \cdots, s^{\mu_1(\mu_2-1)}],$$

$$H = \{1, t, \cdots, t^{\nu_1-1}\}[1, s, \cdots, s^{\mu_1-1}].$$

Now, solving $A_{11}\langle B_{11}\rangle = C_{11}\langle D_{11}\rangle$ by the method of §3, we find

$$A_{11} = a_{11}E, \quad B_{11} = b_{11}F, \quad C_{11} = c_{11}G, \quad D_{11} = d_{11}H,$$

where $a_{11}b_{11}=c_{11}d_{11}$. Similarly solving $A_{xy}\langle B_{zu}\rangle = C_{xy}\langle D_{zu}\rangle$, we have

$$A_{xy} = a_{(x-1)\nu_1+1,(y-1)\mu_1+1}E, \qquad B_{zu} = b_{zu}F,$$

$$C_{xy} = a_{(x-1)\nu_2+1,(y-1)\mu_2+1}G, \qquad D_{zu} = d_{zu}H,$$

where $a_{(x-1)\nu_1+1,(y-1)\mu_1+1}b_{zu} = c_{(x-1)\nu_2+1,(y-1)\mu_2+1}d_{zu}$ for all values of x, y, z, u. Hence

$$\frac{a_{(x-1)\nu_1+1,(y-1)\mu_1+1}}{c_{(x-1)\nu_2+1,(y-1)\mu_2+1}} = \frac{d_{zu}}{b_{zu}} = q,$$

where q is a constant for all values of x, y, z, u. We notice that E, F, G, H are the same for all values of x, y, z, u, for otherwise we would find some B_{zu} having two different values at once. It follows that

$$A = \begin{bmatrix} A_{11}, & \cdots, A_{1k_1} \\ \cdots & \cdots \\ A_{k_11}, & \cdots, A_{k_1k_1} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}, & \cdots, a_{1,(k_1-1)\mu_1+1} \\ \cdots & \cdots \\ a_{(k_1-1)\nu_1+1,1} & \cdots a_{(k_1-1)\nu_1+1,(k_1-1)\mu_1+1} \end{bmatrix}\langle E\rangle$$

$$= q \begin{bmatrix} c_{11}, & \cdots, & c_{1,(k_1-1)\mu_2+1} \\ \cdots & \cdots & \cdots \\ c_{(k_1-1)\nu_2+1,1}, & \cdots, & c_{(k_1-1)\nu_2+1,(k_1-1)\mu_2+1} \end{bmatrix} \langle E \rangle .$$

Similarly

$$C = \begin{bmatrix} c_{11}, & \cdots, & c_{1,(k_1-1)\mu_2+1} \\ \cdots & \cdots & \cdots \\ c_{(k_1-1)\nu_2+1,1}, & \cdots, & c_{(k_1-1)\nu_2+1,(k_1-1)\mu_2+1} \end{bmatrix} \langle G \rangle .$$

In the same way it can be shown that

$$B = F\left\langle \begin{bmatrix} b_{11}, & \cdots, & b_{1k_2} \\ \cdots & \cdots & \cdots \\ b_{k_21}, & \cdots, & b_{k_2k_2} \end{bmatrix} \right\rangle \text{ and } D = qH\left\langle \begin{bmatrix} b_{11}, & \cdots, & b_{1k_2} \\ \cdots & \cdots & \cdots \\ b_{k_21}, & \cdots, & b_{k_2k_2} \end{bmatrix} \right\rangle .$$

The above values for A, B, C, D give the most general solution of the equation

$$A\langle B \rangle = C\langle D \rangle .$$

THE UNIVERSITY OF EDINBURGH

A PLANE ELLIPTIC CURVE OF ORDER $4k+2$, WITH SINGULARITIES ALL REAL AND DISTINCT, AND AUTOPOLAR BY $4k+4$ CONICS

BY D. C. DUNCAN

Comparatively few curves are known which are their own polar reciprocals (that is, autopolar) by a finite number of conic sections. The so-called W-curves of Klein and Lie are autopolar by a single infinity of conics, and these are the only curves having this property,* but these rational curves are highly specialized in that all their point and line singularities are concentrated at two places of the loci. Appell† has shown how to find all curves which are autopolar by one given conic. Haskell‡ has shown how Appell's method can be used to derive curves which are autopolar by a *finite* number of conics. This latter device, though sound in principle, involves severe algebraic difficulties in the very magnitude of the degree of the equations to be treated. Consequently only a very few curves whose singular elements are distinct are known to be autopolar by more than one conic. The quartic§ with two cusps and a node is autopolar by two conics. The quintic with five cusps is autopolar by one, two, or six conics, the number of polarizing conics depending upon the symmetry of the quintic. The completely‖ symmetric quintic with three cusps and three nodes is autopolar by four conics; if the nodes and cusps unite to form three rhamphoid cusps, the quintic is still autopolar by three conics. The rational¶ quintic of class five in general is not autopolar. Beyond the fifth order only a few** isolated special cases of autopolar curves have been published. In the present and succeeding papers the writer proposes to show that there exist rational and

* Fouret, Bulletin de la Société Philomathique, 1877.

† Appell, Nouvelles Annales, (3), vol. 13 (1894).

‡ Haskell, Proceedings of the Toronto Congress, vol. I, 1924, pp. 715–717.

§ Wear, American Journal of Mathematics, vol. 42, pp. 97–118; note also American Journal of Mathematics, vol. 51, pp. 482–490.

‖ Duncan, University of California Publications, 1928.

¶ Swinford, University of California Publications, 1929.

** Duncan, this Bulletin, vol. 39 (1933), pp. 589–592.

elliptic curves of every order above five having their singular elements all distinct and which are autopolar by a finite number of conics, the number of polarities actually increasing with the order of the special types considered. Moreover these curves can be sketched approximately with astonishing ease as the sequel will show. In the present paper only elliptic curves of order $4k+2$ are treated. The equations and polarities are exhibited for the two curves of lowest degree, namely six and ten.

The following discussion indicates that

$$\rho^{2(2k+1)} \sin^2 (2k+1)\theta - (\rho^2 - 1)^3 \prod_{i=1}^{k-1} (\rho^2 - \alpha_i^2)^2 = 0, \tag{1}$$

$$(1 < \alpha_1 < \alpha_2 < \cdots < \alpha_{k-1}),$$

is the equation in polar coordinates of a curve of order $4k+2$, elliptic, with singularities all distinct, and autopolar by $4k+4$ conics. The constants α_i^2 are the unique set of positive constants a_i which cause the equation

$$x^{2k+1} - (x-1)^3 \prod_{i=1}^{k-1} (x - a_i)^2 = 0 \tag{2}$$

to have $k-1$ double roots. The uniqueness of the set of a_i is shown below. One obtains (2) from (1) by setting

$$\rho^2 = x, \qquad \alpha_i^2 = a_i, \qquad \theta = \pi/(4k+2).$$

Observe that (1) represents a locus having the following properties.

(i) It is invariant under rotations about the origin through the angles $\lambda\pi/(2k+1)$, $(\lambda = 1, 2, \cdots, 4k+1)$.

(ii) On the line $\theta = 0$ it has cusps at $\rho = \pm 1$ and crunodes at $\rho = \pm\alpha_i$.

(iii) On the line $\theta = \pi/(4k+2)$ it has a biflecnode at infinity and $2(k-1)$ crunodes corresponding to square roots of a_i in (2).

(iv) The locus has $4k+2$ cusps and $(2k+1)(4k-3)$ nodes, the requisite number for an elliptic self-dual curve of order $4k+2$.

(v) The nodes at infinity are biflecnodes whose inflexional tangents are perpendicular to cuspidal tangents.

The uniqueness of the a_i in (2) is to be noted geometrically as follows. In $y = x^{2k+1} - (x-1)^3 \Pi_{i=1}^{k-1} (x-a_i)^2$, let $y_1 = x^{2k+1}$, $y_2 = (x-1)^3 \Pi_{i=1}^{k-1} (x-a_i)^2$, $y^{(1)} = (x-a_1)^2$, $y^{(2)} = (x-a_2)^2$, $\cdots$,

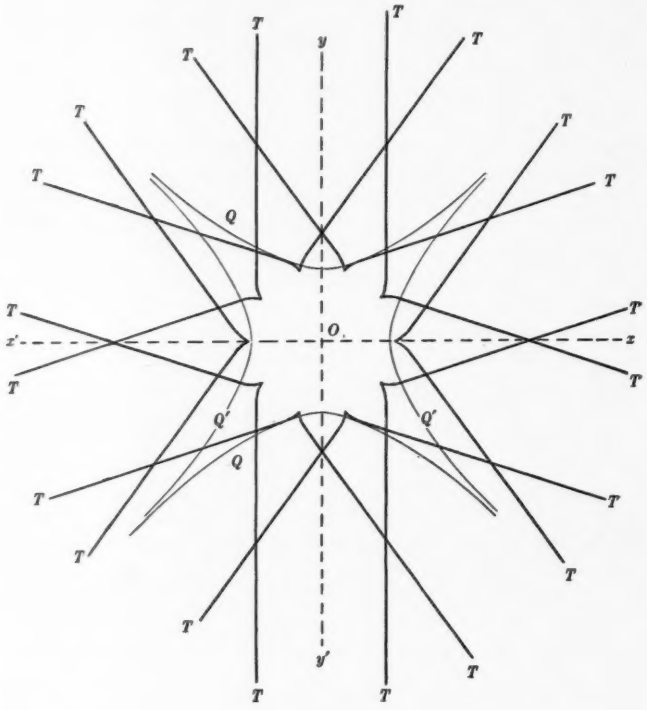
$y^{(k-1)}=(x-a_{k-1})^2$. Now for $1<a_1<a_2\cdots<a_{k-1}$, $y^{(i)}$ represent $k-1$ distinct parabolas with vertices on the x-axis and opening upward; hence y_2 represents a curve having ordinary contacts with the x-axis at $x=a_i$ and a contact of second order at $x=1$, hence $k-1$ maxima. Accordingly, for only one position of a_i will y_2 touch y_1 in $k-1$ distinct points. These are the desired unique values for a_i, $(i=1, 2, 3, \cdots, k-1)$. Therefore only one such locus (1) can exist.

Two concentric circles $x^2+y^2\pm r^2z^2=0$ polarize the cusps into the inflexional tangents (asymptotes perpendicular to cuspidal tangents). By the uniqueness of (1) these circles must autopolarize the entire locus. Since only *even* powers of x, y appear in the rectangular Cartesian form of (1), the locus is also autopolar by the rectangular hyperbolas $x^2-y^2\pm r^2z^2=0$. By symmetry there must be $4k+2$ such hyperbolas, or $4k+4$ autopolarizing conics in all.

Non-degeneracy of the locus is observed from the following considerations. The locus has $2k+1$ biflecnodes at infinity, hence no finite inflexions; the nodes are all on cuspidal tangents, or on the bisectors of adjacent cuspidal tangents. For example, on the cuspidal tangent $\theta=0$ there are $k-1$ nodes. They can be produced only by the intersections of arms of cusps on either side of this ray, that is, by half the arms of $2(k-1)$ cusps. These arms must be asymptotic, joining with $2(k-1)$ other cuspidal arms (since no finite inflexions are permitted). Hence the second arms of the cusps adjacent to that at (0, 1) must have a vertical asymptote. Symmetry thus permits the curve to be sketched approximately by drawing secant lines through alternate points of $2(2k+1)$ points equally spaced about the unit circle, omitting the chords within the circle. Accordingly the locus is bipartite, each branch consisting of a set of alternate cusps with their infinitely long arms. If each branch is a proper curve, it must be also self-dual and of order $n/2$, (where $n=4k+2$), and possess $n/2$ cusps, and consequently $n(n-10)/8$ nodes. But one branch approximates $n/2$ secant lines through $n/2$ *consecutive* points equally spaced about the unit circle, and can then have only $n(n-6)/8$ nodes ($\neq n(n-10)/8$, as required). Therefore the locus is non-degenerate and consists of two congruent branches. These branches are polar reciprocal by $2k+2$ of the polarities, and individually autopolar by the other $2k+2$ polarities. More-

over the locus is invariant under $4k$ correlations in addition to the $4k+4$ polarities and $2(4k+2)$ collineations.

The equations and polarities are next exhibited for the two curves of lowest degree of this set.



T: Bipartite 10-ic autopolar by 10 rectangular hyperbolas and 2 circles.

Q, Q': Two of the 5 sets of conjugate hyperbolas which autopolarize the locus.

Q autopolarizes the two branches individually and has 4 real contacts with T.

Q' interchanges the two branches by the polarity.

CASE I.

$$k = 1, \qquad (n = 6).$$

Equation: $\rho^6 \sin^2 3\theta - (\rho^2-1)^3 = 0$, or $y^2(3x^2-y^2)^2 - (x^2+y^2-z^2)^3 = 0$. Cusps: $(\pm 1, 0, 1)$, $(\pm 3^{1/2}, \pm 1, 2)$. Biflecnodes: $(0, 1, 0)$, $(2, \pm 3^{1/2}, 0)$. The eight polarities are readily noted among the twelve correlations:

$$\begin{array}{ccccc} u & : & v & : & w \\ \pm 3^{1/2}x & : & \pm 3^{1/2}y & : & z \\ 3^{1/2}x \pm 3y & : & 3x \pm 3^{1/2}y & : & \pm 2z \end{array}$$

CASE II. $$k = 2, \qquad (n = 10).$$

Equation:

$$\rho^{10} \sin^2 5\theta - (\rho^2-1)^3(\rho^2-\alpha^2)^2 = 0,$$

$$\alpha^2 = 5[10(1/10)^{1/3} + 3 + 2\cdot 10^{1/3}]/9,$$

or

$$y^2(5x^4 - 10x^2y^2 + y^4)^2 - (x^2 + y^2 - z^2)^3(x^2 + y^2 - \alpha^2 z^2)^2 = 0,$$

$$(\alpha = 2.576\cdots).$$

Twelve polarities are readily noted among the twenty correlations.

$$u : v : w$$

$$5\left(x\cos\frac{k\pi}{10} - y\sin\frac{k\pi}{10}\right) : \pm 5\left(x\sin\frac{k\pi}{10} + y\cos\frac{k\pi}{10}\right) : (2\alpha^2+3)^{1/2}z,$$

$$(k = 0, 1, \cdots, 9).$$

SUMMARY. A completely symmetric, elliptic curve exists of order $4k+2$, having $4k+2$ cusps and $(2k+1)(4k-3)$ nodes, all distinct. The locus is invariant under $2(4k+2)$ correlations of which $4k+4$ are polarities. The locus is very approximately realized by drawing secant lines through the alternate points of $4k+2$ points equally spaced about the unit circle, omiting the chords. The locus is bipartite. Its $(2k+1)(4k-3)$ bitangents are very nearly the secants joining all pairs of the points of the circle which are separated by at least two of the points and which are not the ends of the same diameter.

THE UNIVERSITY OF CALIFORNIA

AN INVERSE PROBLEM IN DIFFERENTIAL EQUATIONS*

BY R. E. LANGER

1. *Introduction.* The differential equation as a tool requires no introduction to either the mathematician or the applied scientist. Problems in endless variety are continually solved through this medium, the process almost invariably beginning with an epitome of the problem's essential characteristics in the form of a differential equation, which is thus determined explicitly both as to its structure and its coefficients, and proceeding thence to a deduction of the form or properties of a suitable solving function.

The present note is devoted to a problem in which this customary order of events is in large measure reversed. The formulation of the problem yields in this case the structural form of a differential equation, and beyond this the existence of a solution which satisfies certain specified conditions. From these data the determination of the equation itself, that is, of its coefficient function, is required and constitutes the solution of the problem.

2. *The Physical Problem.*† In the investigation of shallow geological structures, and in the study of the electrical resistivity of the earth's crust at depths below the surface, an appropriate experimental procedure centers around the supply of a direct electric current through a small electrode to the surface of the earth. The electrical potentials which result at the surface of the earth are measurable at all distances from the electrode, and constitute entirely the immediately obtainable data. From them it is desired to compute, if possible, the conductivity of the earth below as a function of the depth.

With the idealizations involved in regarding the conductivity as a differentiable point function depending upon the depth alone, and in taking the ground as a horizontally uniform in-

* Presented to the Society, April 14, 1933.

† The problem and its formulation as outlined in this section are due to Professor L. B. Slichter of the Massachusetts Institute of Technology. A complete and detailed discussion of it is given in a geophysical paper by L. B. Slichter, in Physics, vol. 4, Sept., 1933.

finite half-space, the problem may be formulated in the following way. Let x, the depth, and ρ, the horizontal distance from the electrode, be taken as cylindrical coordinates with origin at the electrode. The electrical potential $\phi(\rho, x)$ may be shown then to satisfy the differential equation

$$\sigma(x)\left[\frac{\partial^2\phi}{\partial\rho^2}+\frac{1}{\rho}\frac{\partial\phi}{\partial\rho}+\frac{\partial^2\phi}{\partial x^2}\right]+\frac{d\sigma(x)}{dx}\cdot\frac{\partial\phi}{\partial x}=0,$$

in which $\sigma(x)$ denotes the earth's conductivity. The substitution

$$\phi(\rho, x) = U(\rho)u(x)$$

separates the variables and resolves the equation into the components

$$U'' + \rho^{-1}U' + \lambda^2 U = 0,$$
$$\{\sigma u'\}' - \lambda^2\sigma u = 0.$$

The first of these is a Bessel equation, and, since the potential is to remain finite and vanish at infinity like the reciprocal of the distance, it must be concluded that $U(\rho)\equiv J_0(\lambda\rho)$.

The second component equation is of the Sturm-Liouville type, and since $\sigma(x)$ is positive, its solutions are of exponential form. Let $u_1(x, \lambda)$ denote a solution which is positive, and, as a function of x, monotonically decreasing. The condition that $\partial\phi/\partial x$ vanish everywhere at the surface except at the electrode leads by familiar reasoning to the formula

$$\phi(\rho, x) = \frac{-c}{2\pi a\sigma(0)}\int_0^\infty \frac{u_1(x, \lambda)}{u'(0, \lambda)}J_0(\lambda\rho)\sin\lambda a d\lambda,$$

in which c is the current, and a the radius of the electrode. In terms of the abbreviation

$$\Omega(\lambda) \equiv -\lambda u_1(0, \lambda)/u_1'(0, \lambda),$$

the surface potentials are accordingly given by

$$\phi(\rho, 0) = \frac{c}{2\pi\sigma(0)}\int_0^\infty \Omega(\lambda)\frac{\sin\lambda a}{\lambda a}J_0(\lambda\rho)d\lambda,$$

a relation which by the Fourier-Bessel integral theorem may be inverted into the form

$$\Omega(\lambda)\left[c\,(\sin \lambda a)/(2\pi a)\lambda^2\sigma(0)\right] = \int_0^\infty \phi(\rho, 0)J_0(\lambda\rho)\rho d\rho.$$

The formulas thus derived reveal the significant fact that the functions $\Omega(\lambda)$ and $\phi(\rho, 0)$ are each uniquely determined by the other, and hence that the information embodied in the surface potential data is completely embraced in the function $\Omega(\lambda)$ when the latter is given for $0<\lambda<\infty$. It is to be shown how from these data the conductivity function $\sigma(x)$ may be computed.

3. *The Mathematical Problem.* The considerations sketched in brief above may be looked upon as having crystallized the physical problem into the following somewhat idealized mathematical one.

On some interval $0 \leqq x \leqq h$ a certain function $\sigma(x)$ is known to be analytic and positive. Beyond this it is known that the differential equation

$$\frac{d}{dx}\left\{\sigma(x)\frac{du}{dx}\right\} - \lambda^2\sigma(x)u = 0 \tag{1}$$

possesses a solution $u_1(x, \lambda)$ which has the properties:

(i) that for λ on the range $(0, \infty)$ and x on the interval $(0, h)$ the relations

$$u_1(x, \lambda) > 0,\ u_1'(x, \lambda) < 0, \tag{2}$$

are satisfied; while

(ii) at $x=0$, the boundary condition

$$\Omega(\lambda)u_1'(0, \lambda) + \lambda u_1(0, \lambda) = 0 \tag{3}$$

is fulfilled, $\Omega(\lambda)$ being a function which is compatible with the preceding hypothesis, and which is known and given for $0<\lambda<\infty$. The function $\sigma(x)$ is to be computed.

A process for the desired computation may be deduced as follows. Let the equation (1) be written in the form

$$u''(x, \lambda) + \frac{\sigma'(x)}{\sigma(x)}u'(x, \lambda) - \lambda^2 u(x, \lambda) = 0. \tag{1a}$$

It is then readily seen to possess a fundamental set of solutions which are represented asymptotically (as to λ) by a pair of expressions

$$e^{\pm\lambda x}S(x, \pm\lambda),$$

in which

$$S(x, \lambda) = \sigma^{-1/2}(x)\left[1 + \sum_{n=1}^{\infty} \alpha_n(x)\lambda^{-n}\right].$$

Hence the solution which satisfies the condition (3), and includes a suitable factor independent of x, is representable by a formula

$$u_1(x, \lambda) \sim e^{-\lambda x}S(x, -\lambda) + \frac{\Omega(\lambda) - 1}{\Omega(\lambda) + 1}e^{\lambda x}S(x, \lambda).$$

Now unless the first term on the right of this expression is dominant for all values of x on $(0, h)$, the one or the other of the conditions (2) will inevitably be violated when λ is sufficiently large. It follows, therefore, that

$$u_1(x, \lambda) \sim e^{-\lambda x}S(x, -\lambda), \tag{4}$$

a form which on substitution into the condition (3) is found to impose upon the function $\Omega(\lambda)$ a condition of compatibility to the effect that it admit of a representation

$$\Omega(\lambda) \sim 1 + \sum_{n=1}^{\infty} \omega_n\lambda^{-n}. \tag{5}$$

Since by hypothesis the function $\Omega(\lambda)$ is given, the infinite set of constants ω_n, $(n=1, 2, 3, \cdots)$, is to be considered as known.

Let the function $v(x, \lambda)$ be defined by the formula

$$v(x, \lambda) = -\lambda u_1(x, \lambda)/u_1'(x, \lambda). \tag{6}$$

It is then found on the one hand in virtue of (1a) to satisfy the Riccati equation

$$v'(x, \lambda) - \frac{\sigma'(x)}{\sigma(x)}v(x, \lambda) - \lambda v^2(x, \lambda) + \lambda = 0, \tag{7}$$

and on the other hand, in virtue of (3), (4) and (5), to be asymptotically representable in a form

$$v(x, \lambda) \sim 1 + \sum_{n=1}^{\infty} v_n(x)\lambda^{-n}, \tag{8}$$

the coefficients $v_n(x)$ being analytic on $(0, h)$, and satisfying the boundary relations

$$v_n(0) = \omega_n, \qquad (n = 1, 2, 3, \cdots). \tag{9}$$

A relation

$$v^{-1}(x) \sim 1 + \sum_{n=1}^{\infty} r_n(x)\lambda^{-n} \tag{8a}$$

then follows, the coefficients being related to those of the series (8) by the recurrence formulas*

$$r_n(x) = -\sum_{j=1}^{n-1} v_j(x) r_{n-j}(x) - v_n(x), \quad (n = 1, 2, 3, \cdots). \tag{10}$$

If the equation (7) is now written in the form

$$\frac{\sigma'(x)}{\sigma(x)} = -\lambda v(x, \lambda) + \frac{v'(x, \lambda) + \lambda}{v(x, \lambda)},$$

and the series (8) and (8a) are formally substituted in it, it is found as a result that

$$\frac{\sigma'(x)}{\sigma(x)} \sim -v_1(x) + r_1(x) + \sum_{n=1}^{\infty} \lambda^{-n} \left[v_n' + r_{n+1} - v_{n+1} + \sum_{j=1}^{n-1} v_j' r_{n-j} \right].$$

The function $\sigma(x)$, however, does not depend upon λ. Hence it must be concluded that

$$\frac{\sigma'(x)}{\sigma(x)} = -v_1(x) + r_1(x), \tag{11}$$

and

$$\sum_{j=1}^{n-1} v_j' r_{n-j} + v_n' = v_{n+1} - r_{n+1}, \quad (n = 1, 2, 3, \cdots). \tag{12}$$

Of the equations (12) the first p in number when taken together constitute a linear algebraic system for the unknowns $v_1', v_2', \cdots, v_p'$. The system has a determinant of value unity and on solution yields the formulas

* In this as well as in subsequent formulas an indicated sum is to be considered as zero if the upper index of summation is less than the lower one.

$$(13)\quad v_p'(x) = \begin{vmatrix} 1 & 0 & \cdots 0 & v_2 - r_2 \\ r_1 & 1 & \cdots 0 & v_3 - r_3 \\ r_2 & r_1 & \cdots 0 & v_4 - r_4 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ r_{p-2} & r_{p-3} \cdots 1 & v_p - r_p \\ r_{p-1} & r_{p-2} \cdots r_1 & v_{p+1} - r_{p+1} \end{vmatrix}, \quad (p = 1, 2, 3, \cdots).$$

These formulas may be materially simplified as follows. The iteration of the formula (10) yields the relations

$$-r_n = \sum_{j=2}^{n-1}(v_j - v_1 v_{j-1}) r_{n-j} + (v_n - v_1 v_{n-1}),$$

from which it may be seen that if to the last column in the determinant (13) there is added the combination

$$-\sum_{j=1}^{p-1}(v_{j+1} - v_1 v_j) \text{ times the } j\text{th column},$$

the effect is formally to replace the elements $r_2, r_3, \cdots, r_p$ by zeros and in the case of the last row to replace $-r_{p+1}$ by $v_{p+1} - v_1 v_p$. In precisely the same way it will be seen upon reference to the formulas (10), that if to the first column in (13) there is added the combination

$$\sum_{j=2}^{p-1} v_{j-1} \text{ times the } j\text{th column},$$

the formal effect is to replace the elements $r_1, r_2, \cdots, r_{p-2}$ by zeros, and in the case of the last row to replace r_{p-1} by $-v_{p-1}$. Similar reductions of the remaining columns may likewise be made, the formula (13) being reduced in consequence to the form

$$v_p'(x) = \begin{vmatrix} 1 & 0 & \cdots & 0 & v_2 \\ 0 & 1 & \cdots & 0 & v_3 \\ 0 & 0 & \cdots & 0 & v_4 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & \cdots & 1 & v_p \\ -v_{p-1} & -v_{p-2} \cdots & -v_1 & 2v_{p+1} - v_1 v_p \end{vmatrix}.$$

If in this, finally, there is added to the last row the combination

$$\sum_{j=1}^{p-1} v_{p-j} \text{ times the } j\text{th row},$$

the determinant disintegrates into the formulas

$$\begin{aligned} v_1'(x) &= f_1(v_1, v_2), \\ v_n'(x) &= \phi_n(v), \qquad (n = 2, 3, 4, \cdots), \end{aligned} \tag{14}$$

in which

$$\begin{aligned} f_1(v_1, v_2) &= 2v_2(x) - v_1^2(x), \\ \psi_n(v) &= 2v_{n+1}(x) + \sum_{j=1}^{n-2} v_{j+1}(x)v_{n-j}(x). \end{aligned} \tag{15}$$

Let the functions $f_2, f_3, \cdots$, be defined successively by the recurrence formula

$$f_n(v_1, v_2, \cdots, v_{n+1}) = \sum_{i=1}^{n} \phi_i(v) \frac{\partial f_{n-1}}{\partial v_i}. \tag{16}$$

Then it is readily seen that repeated differentiation of the first of formulas (14), and substitution from the remaining formulas, gives the expressions

$$\frac{d^n v_1(x)}{dx^n} = f_n(v_1, v_2, \cdots, v_{n+1}), \quad (n = 1, 2, 3, \cdots), \tag{17}$$

for the derivatives of the function $v_1(x)$.

For general values of x the functions involved in the right member of (17) are not known. They are given, however, for $x=0$ by (9), whence

$$\left.\frac{d^n v_1}{dx}\right]_{x=0} = f_n(\omega_1, \omega_2, \cdots, \omega_{n+1}), \quad (n = 1, 2, 3, \cdots). \tag{18}$$

With these values available the MacLaurin expansion of the function $v_1(x)$ is computable, and since the first of the formulas (10) gives to (11) the form

$$\frac{\sigma'(x)}{\sigma(x)} = -2v_1(x),$$

the desired computation of the function $\sigma(x)$ has been accomplished.

THE UNIVERSITY OF WISCONSIN

ON THE SUMMABILITY AND GENERALIZED SUM OF A SERIES OF LEGENDRE POLYNOMIALS*

BY W. C. BRENKE

1. *Introduction.* The results obtained in this paper are as follows.

(A) *The series of Legendre polynomials* $\sum n^p X_n(x)$, *where* p *is a positive integer, is summable* (H, p) *for* $-1<x<1$, *and summable* $(H, p+1)$ *for* $-1\leqq x<1$.

(B) *The generalized sum over the range* $-1<x<1$ *is*

$$\sum_1^\infty n^p X_n(x) = -\frac{1}{2(2y)^{p-1/2}}\begin{vmatrix} 2 & 0 & 0 & 0 & \cdots & 0 & 1 \\ y & 2y & 0 & 0 & \cdots & 0 & 1 \\ A_2^3 & y & 2y & 0 & \cdots & 0 & 1 \\ A_3^4 & A_2^4 & y & 2y & \cdots & 0 & 1 \\ \cdot & \cdot & & & & \cdot & \cdot \\ \cdot & \cdot & & & & \cdot & \cdot \\ A_{p-2}^{p-1} & A_{p-3}^{p-1} & \cdots & y & & 2y & 1 \\ A_{p-1}^{p} & A_{p-2}^{p} & \cdots & A_2^p & & y & 1 \end{vmatrix}$$

where

$$y = 1 - x; \qquad A_i^p = {}_pC_i + (-1)^i {}_{p-1}C_i; \qquad (p > 2).$$

2. *The Cases* $p=0, 1, 2$. We first obtain these results for $p=0$, 1, 2. Let p be a positive integer, $S_{n,p}$ the sum of the first n terms of the series $\sum n^p X_n(x)$, $S_{n,p}^{(p)}$ the pth Hölder mean of $S_{n,p}$, and $S^{(p)}$ the limit of this mean for $n\to\infty$.

The generating function of the Legendre polynomials gives us at once the sum of the convergent series

$$(1) \qquad \sum_1^\infty X_n(x) = S^{(0)} = [2(1-x)]^{-1/2} - 1, \qquad (-1 < x < 1).$$

We can readily find $S^{(1)}$ by use of the recursion formula

$$(2) \qquad (2m+1)xX_m = (m+1)X_{m+1} + mX_{m-1},$$

* Presented to the Society, November 26, 1932.

which we write in the form

$$2mxX_m = (m+1)X_{m+1} + (m-1)X_{m-1} + X_{m-1} - xX_m.$$

Summing from $m=1$ to $m=n$ we obtain, after some reductions,

$$2(x-1)S_{n,1} = (1-x)S_{n,0} + X_0 - X_1 + (n+1)X_{n+1} - nX_n - X_n.$$

The first mean now gives

$$2(x-1)S^{(1)}_{n,1} = (1-x)S^{(1)}_{n,0} + (1-x) + [(n+1)X_{n+1} - X_1 - S_{n,0}]/n.$$

For $n\to\infty$ we then have

$$(1') \qquad 2(x-1)S^{(1)} = (1-x)S^{(0)} + (1-x).$$

To obtain $S^{(2)}$ we multiply (2) by m and arrange in the form

$$2m^2xX_m = u(m+1)X_{m+1} + v(m-1)X_{m-1} - mxX_m,$$

where u and v are quadratic functions of the indicated arguments. Summing and taking two successive mean values we obtain after some reductions

$$2(x-1)S^{(2)}_{n,2} = (1-x)S^{(2)}_{n,1} + S^{(2)}_{n,0} + 1 - (1/n)\sum_{r=1}^{n}[3S_{r,1} - r(r+1)X_{r+1} + S_{r,0}]/r.$$

The last term vanishes for $n\to\infty$. We note also that the existence of the limit of any mean ensures the existence of the same limit for the higher means. Hence for $n\to\infty$ the preceding equation becomes

$$(1'') \qquad 2(x-1)S^{(2)} = (1-x)S^{(1)} + S^{(0)} + 1.$$

3. *Proof of* (A) *and* (B). We shall now show that

$$(3)\quad 2(x-1)S^{(p)} = (1-x)S^{(p-1)} + \sum_{t=2}^{p-1} A^p_t S^{(p-t)} + S^{(0)} + 1, (p>2);$$

provided that $S^{(r)}$ exists for $r=0, 1, 2, \cdots, p-1$, and A^p_t, expressed in terms of binomial coefficients, is

$$A^p_t = {}_pC_t + (-1)^t {}_{p-1}C_t.$$

Assuming the existence of the means of order less than p we first prove the following lemma.

LEMMA. *The means of order p of n^pX_n and of $(n+1)^pX_{n+1}$ have the limit zero for $n\to\infty$, when $-1<x<1$.*

To prove this we multiply (2) by m^{p-1}, express the coefficients of X_{m+1} and X_{m-1} on the right in powers of $m+1$ and $m-1$, respectively, sum from $m=1$ to $m=n$, and write the result in the form

$$(4)\quad 2(x-1)S_{n,p} = (1-x)S_{n,p-1} + \sum_{i=2}^{p-1} A_i^p S_{n,p-i} + S_{n,0} + X_0 + n^{p-1}(n+1)X_{n+1} - (n+1)^pX_n.$$

Successive application of this formula enables us to write

$$2(x-1)S_{n,p} = w_n + u_p(n+1)X_{n+1} + v_p(n)X_n,$$

where $w_n = A_pS_{n,0} + B_pX_0$, A_p and B_p being independent of n, and $u_p(n+1)$, $v_p(n)$ are polynomials of degree p in the indicated arguments. Denoting the mean of order k of n^pX_n by $M_{n,p}^{(k)}$, we have

$$M_{n,p}^{(1)} = S_{n,p}/n,$$

which, by use of the preceding equation, we may write in the form

$$2(x-1)M_{n,p}^{(1)} = w_n^{(1)} + u_{p-1}(n+1)X_{n+1} + v_{p-1}(n)X_n,$$

where $w_n^{(1)} = (w_n + C_pX_{n+1} + D_pX_n)/n$, C_p and D_p being independent of n, and where the coefficients of the last two terms are polynomials of the indicated degree and arguments. We note that $w_n^{(1)}$ and all of its means$\to 0$ when $n\to\infty$. Proceeding in this way we have

$$2(x-1)M_{n,p}^{(2)} = w_n^{(2)} + u_{p-2}(n+1)X_{n+1} + v_{p-2}(n)\,X_n,$$

and finally

$$2(x-1)M_{n,p}^{(p)} = w_n^{(p)} + u_0(n+1)X_{n+1} + v_0(n)X_n,$$

where each $w_n^{(r)}$ and its means$\to 0$ when $n\to\infty$. The means of $(n+1)^pX_{n+1}$ may be treated similarly.

COROLLARY 1. *The mean of order p of $u_p(n+1)X_{n+1}$ and of $v_p(n)X_n$ approach zero when $n\to\infty$, for $-1<x<1$.*

COROLLARY 2. *For $x=-1$ the mean of order $p+1$ of each of the preceding expressions vanishes when $n\to\infty$. This follows from $X_n(-1)=(-1)^n$.*

Now (3) is obtained at once by taking the limit of the mean of order p of (4). But the values of $S^{(0)}$, $S^{(1)}$, $S^{(2)}$ already found show that (3) holds for $p=3$. Hence it holds for positive integral values of $p>2$.

The result under (B) is obtained by expressing each $S^{(r)}$, $(r=1, 2, \cdots, p)$, in terms of the sums of lower order by use of (1′), (1″), (3) and solving this system of equations for $S^{(p)}$.

When $x=-1$, Corollary 2 shows that the series $\sum(-1)^n n^p$ is summable $(H, p+1)$ and a new form is obtained for its sum by putting $y=2$ in the formula under (B).

THE UNIVERSITY OF NEBRASKA

ERRATA

This Bulletin, volume 37 (1931), pages 759–765:

On page 759, line 12, *for* $P_1\equiv P, P_2, \cdots$ *read* $P\equiv P_1, P_2, \cdots$.

On page 764, line 9, second parenthesis, *for* (x_1x_2, x_2, x_1x_4) *read* (x_1x_2, x_2^2, x_1x_4).

On page 764, line 8 from the bottom, second parenthesis, *for* $(z_2, z_3, \epsilon_3 z_4)$ *read* $(\epsilon z_2, z_3, \epsilon^3 z_4)$.

W. R. HUTCHERSON

This Bulletin, volume 39 (1933), p. 589:

Lines 13–14, omit the words: *one point of inflexion*; and add the sentence: *A point of inflexion lies at infinity on each bisector of the angles formed by adjacent cuspidal tangents.*

D. C. DUNCAN

OFFICIAL COMMUNICATIONS

Meetings of the Society have been fixed at the following times and places:

NEW YORK CITY, October 28, 1933.

Abstracts must be in the hands of Associate Secretary J. R. Kline, 501 West 116th St., New York City, not later than October 7. Professor R. L. Jeffery will deliver an address on *Theories of integration*. Abstracts received by the Secretaries by September 9 will appear in the September issue of the BULLETIN.

CINCINNATI, OHIO, December 1-2, 1933.

Abstracts must be in the hands of Associate Secretary M. H. Ingraham, University of Wisconsin, Madison, Wis., not later than November 11. Abstracts received by the Secretaries by November 4 will appear in the November issue of the BULLETIN. A joint meeting regarding spinor analysis will be held with the Physical Society; Professor Oswald Veblen will represent the Society at this meeting; and Professor Otto La Porte will speak on *Spinor analysis*. Professor R. D. Carmichael will deliver an address entitled *Functions of exponential type*.

PASADENA, CALIFORNIA, December 2, 1933.

Abstracts must be in the hands of Associate Secretary T. M. Putnam, University of California, Berkeley, Calif., not later than November 11. Abstracts received by the Secretaries by November 4 will appear in the November issue of the BULLETIN.

CAMBRIDGE, MASSACHUSETTS, ANNUAL MEETING, December 27-29, 1933.

This meeting is to be held in conjunction with the A.A.A.S. and the M.A.A. There will be a symposium on *General Analysis*. The Bôcher Prize will be awarded. Abstracts must be in the hands of the Secretary of the Society, 501 West 116th St., New York City, not later than November 30, 1933. Abstracts received by the Secretaries by November 4 will be published in the November issue of the BULLETIN.

R. G. D. RICHARDSON, *Secretary of the Society*.

Articles for insertion in the BULLETIN should be addressed to E. R. HEDRICK, Editor of the BULLETIN, University of California at Los Angeles. Reviews should be sent to W. R. LONGLEY, Yale University, New Haven, Conn. Notes should be sent to H. W. KUHN, Ohio State University, Columbus, Ohio.

Subscriptions to the BULLETIN, orders for back numbers, and inquiries in regard to non-delivery of current numbers should be addressed to the American Mathematical Society, 450-459 Ahnaip St., Menasha, Wis., or 501 West 116th St., New York.

The initiation fees and the annual dues of members of the Society (see this BULLETIN, p. 322, May, 1930; and the List of Officers and Members, September, 1932, p. 66), are payable to the Treasurer of the Society, Professor G. W. Mullins, 501 West 116th St., New York City.

WHOLE NUMBER 409

CONTENTS

PAGE

Quantum Mechanics and Asymptotic Series. By G. D. BIRKHOFF 681
Recent Progress on Waring's Theorem and its Generalizations. By L. E. DICKSON 701
Abstract Ideal Theory. By OYSTEIN ORE 728
Normal Division Algebras over Algebraic Number Fields not of Finite Degree. By A. A. ALBERT 746
Further Mean-Value Theorems. By MORRIS MARDEN .. 750
On the Number of Apparent Double Points on a Certain V_s^n in an S_{2s+1}. By B. C. WONG 755
Quadratic Residues in Factorization. By MARSHALL HALL 758
A Complete Census of 4×4 Magic Squares. By D. N. LEHMER 764
On the Theory of Fourier Transforms. By EINAR HILLE and J. D. TAMARKIN 768
Note on the Location of the Critical Points of Green's Function. By J. L. WALSH 775
Simplification of the Set of Four Postulates for Boolean Algebras in Terms of Rejection. By B. A. BERNSTEIN 783
On Section A of *Principia Mathematica*. By B. A. BERNSTEIN 788
On Natural Families of Curves. By C. H. ROWE 793
On the Condition that two Zehfuss Matrices be Equal. By D. E. RUTHERFORD 801
A Plane Elliptic Curve of Order $4k+2$, with Singularities all Real and Distinct, and Autopolar by $4k+4$ Conics. By D. C. DUNCAN 809
An Inverse Problem in Differential Equations. By R. E. LANGER 814
On the Summability and Generalized Sum of a Series of Legendre Polynomials. By W. C. BRENKE 821
Errata 824

For official Communications and Notices, see the inside of the back cover.

GEORGE BANTA PUBLISHING COMPANY, MENASHA, WISCONSIN